



**MINISTÉRIO DA EDUCAÇÃO**  
**UNIVERSIDADE FEDERAL DA GRANDE DOURADOS**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA GERAL /**  
**BIOPROSPECÇÃO**

**ANDREA RENATA DA SILVA ROMERO**

**ESTUDOS GENÔMICOS APLICADOS AO MELHORAMENTO GENÉTICO**  
**DE BOVINOS**

**DOURADOS – MS**

**2017**

**UNIVERSIDADE FEDERAL DA GRANDE DOURADOS FACULDADE DE  
CIÊNCIAS BIOLÓGICAS E AMBIENTAIS**

**ESTUDOS GENÔMICOS APLICADOS AO MELHORAMENTO GENÉTICO DE  
BOVINOS**

Documento apresentado ao Curso de Mestrado do Programa de Pós-Graduação em Biologia Geral/Bioprospecção da Faculdade de Ciências Biológicas e Ambientais da Universidade Federal da Grande Dourados, por Andrea Renata da Silva Romero, para obtenção do título de Mestre em Biologia Geral/Bioprospecção, sob orientação da Prof<sup>a</sup> Dr<sup>a</sup> Alexéia Barufatti Grisolia e co-orientação da Dr<sup>a</sup> Fabiane Siqueira.

**DOURADOS – MS**

**2017**

**Dados Internacionais de Catalogação na Publicação (CIP).**

R763e Romero, Andrea Renata Da Silva  
ESTUDOS GENÔMICOS APLICADOS AO MELHORAMENTO  
GENÉTICO DE BOVINOS / Andrea Renata Da Silva Romero -- Dourados:  
UFGD, 2017.  
63f. : il. ; 30 cm.

Orientadora: Alexeia Barufatti Grisolia  
Co-orientadora: Fabiane Siqueira

Dissertação (Mestrado em Biologia Geral/Bioprospecção) - Faculdade de  
Ciências Biológicas e Ambientais, Universidade Federal da Grande Dourados.  
Inclui bibliografia

1. Genômica. 2. Imputação. 3. GWAS. 4. Polimorfismos. 5. Bovinos. I.  
Título.

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

**©Direitos reservados. Permitido a reprodução parcial desde que citada a fonte.**

"ESTUDOS GENÔMICOS APLICADOS AO MELHORAMENTO GENÉTICO DE BOVINOS".

POR

**ANDREA RENATA DA SILVA ROMERO**

DISSERTAÇÃO APRESENTADA À UNIVERSIDADE FEDERAL DA GRANDE DOURADOS (UFGD), COMO PARTE DOS REQUISITOS EXIGIDOS PARA OBTENÇÃO DO TÍTULO DE MESTRE EM BIOLOGIA GERAL - ÁREA DE CONCENTRAÇÃO: "BIOPROSPECÇÃO".



PROF<sup>a</sup>. DR<sup>a</sup>. ALEXEIA BARUFATTI GRISOLIA  
ORIENTADORA – UFGD



PROF. DR. GILBERTO ROMEIRO DE OLIVEIRA MENEZES  
MEMBRO TITULAR – EMBRAPA GADO DE CORTE



PROF. DR. RUSBEL RAUL ASPILCUETA BORQUIS  
MEMBRO TITULAR – UFGD

Aprovada em 22 de fevereiro de 2017.

## **AGRADECIMENTOS**

*Agradeço primeiramente à Deus, por ter me cercado de pessoas boas ao longo da minha caminhada até aqui, que tanto me incentivaram, me ensinaram e nunca me deixaram desistir.*

*Sou grata à minha incrível família, em especial minha mãe Iza e avó Ana, que sempre foram exemplo de força, honestidade e amor. Sempre me ensinaram que com esforço e fé posso alcançar meus objetivos.*

*À meu fantástico namorado André, meu presente de Deus, por todo carinho, companheirismo, amor e por contribuir para que este trabalho pudesse ser concretizado. Agradeço também à minha sogra Lícia e cunhada Amanda, por toda preocupação e carinho.*

*Aos meus grandes amigos, cujo apreço eu não consigo mensurar, por se fazerem presentes em momentos bons e ruins, nunca me esquecerei de nossas reuniões e das boas risadas, Alisson, Camila, Ana Paula, Jéssica C., Jéssica L.*

*Às minhas irmãs (que não são de sangue), mas que são exemplos de coragem, perseverança, fé, alegria, lealdade e amor. Minha admiração por vocês é infindável Viviane, Laís e Rhaíssa.*

*À minha professora, orientadora, amiga e um pouco mãe Alexéia. Muito obrigada por ter acreditado em mim, por ter me dado a oportunidade de trabalhar com a senhora e sua equipe.*

*Aos meus amigos do laboratório de biotecnologia aplicada a produção animal, que são também minha família. Obrigada pelos momentos de descontração, pela alegria que vocês transmitem Bruno, Alexandre, Jéssica Gonçalves, Jéssica Pereira, Allana, Juliana Fernandes, Juliana Spósito e Adrielle.*

*À Dr<sup>a</sup> Fabiane Siqueira e ao Gustavo Garcia Santiago, por toda paciência e por compartilharem comigo o conhecimento que carregam. Muito obrigada por me acolherem e por terem me ensinado tanto.*

*À Embrapa Gado de Corte e à Embrapa Pecuária Sudeste, pela confiança em disponibilizar os dados, com os quais pude trabalhar.*

*À Universidade Federal da Grande Dourados por ter tido a oportunidade de cursar minha graduação e minha pós graduação.*

*À Associação Brasileira dos Criadores de Canchim (ABCCAN), aos produtores e, especialmente, ao médico veterinário Maury Dorta de Souza Junior, responsável técnico pelas Provas Canchim de Avaliação de Desempenho, por terem gentilmente cedido as amostras biológicas dos animais analisados.*

*Ao Programa de Melhoramento Geneplus/Embrapa por disponibilizar as DEPs dos animais para as características fenotípicas avaliadas neste trabalho.*

*À Capes pela concessão da bolsa de pós-graduação e à Embrapa e Fundect pelo financiamento desta pesquisa.*

*À todos que não citei, mas que de alguma forma contribuíram para o meu desenvolvimento acadêmico e possibilitou a realização deste trabalho.*

## SUMÁRIO

LISTA DE ABREVIATURAS.....	6
LISTA DE FIGURAS .....	7
LISTAS DE TABELAS .....	8
RESUMO GERAL .....	9
ABSTRACT .....	10
REVISÃO BIBLIOGRÁFICA.....	11
Melhoramento genético .....	11
Seleção assistida por marcadores moleculares .....	12
Imputação.....	13
GWAS – Estudos de Associação Genômica Ampla.....	15
Parâmetros do Controle de Qualidade dos marcadores .....	16
Equilíbrio de Hardy-Weinberg .....	16
Menor Frequência Alélica .....	17
Parâmetros do Controle de Qualidade das amostras.....	17
Call rate.....	17
Heterozigose .....	17
Estratificação populacional.....	17
Método Estatístico .....	17
Fenótipos.....	19
Características Fenotípicas Avaliadas.....	19
Comprimento de umbigo .....	19
Pelagem.....	21
Circunferência escrotal .....	22
REFERÊNCIAS BIBLIOGRÁFICAS .....	22
OBJETIVOS ESPECÍFICOS .....	28
CAPÍTULO I.....	29

Resumo .....	30
Introdução .....	31
Material e Métodos .....	31
Resultados e Discussão .....	34
Conclusão.....	37
Perspectivas Futuras .....	38
Agradecimentos .....	38
Referências.....	39
CAPÍTULO II.....	41
Resumo .....	42
Introdução .....	43
Material e Métodos .....	44
Resultados .....	48
Discussão .....	50
Conclusão.....	54
Referências Bibliográficas .....	55
Anexo .....	61

## LISTA DE ABREVIATURAS

ABCCAN	Associação Brasileira de Criadores de Canchim
ACR	Alelos imputados corretamente
CCAC	Cluster de canais de cloreto ativados por cálcio
CE	Circunferência Escrotal
CR	Call rate
DEP	Diferença Estimada na Progenie
EHW	Weinberg equilibrium
GGP-HD	GeneSeek Genomic Profiler
GWAS	Genome Wide Association Study
HD	High Density
IBS	Identical by State
LD	Linkage Disequilibrium
MAF	Minor Allele Frequency
MCMC	Markov Chain Monte Carlo
MGI	Mouse Genomics Informatics
PCAD	Provas Canchim de Avaliação de Desempenho
PERC	Genótipos imputados corretamente
QC	Quality Control
QTL	Quantitative Trait Loci
$r$	Correlação de <i>Pearson</i>
SNP	Single Nucleotide Polymorphism
VAE	Variância Aditiva Explicada
3K	Illumina® Bovine3K
7K	Illumina® BovineLD
50K	Illumina BovineSNP50

## LISTA DE FIGURAS

- Figura 1.** Cenário típico de imputação. As barras na fita dupla DNA representam os marcadores (SNP). Os números (genótipos observados) e interrogação (? - genótipos não observados). Genótipos não obtidos no processo de genotipagem (?) no painel de densidade inferior são estimados os genótipos faltantes são determinados (Fonte: Nascimento, 2016)..... 14
- Figura 2.** Mucosa prepucial interna à bainha (aceitável), mucosa prepucial exposta (não aceitável) (Fonte: Embrapa-São Carlos)..... 20
- Figura 3.** Método de avaliação por escore de comprimento de umbigo pelo sistema de avaliação de bovinos de 1 a 6 (Fonte: PPGZOO UFVJM, 2013). ..... 21
- Figura 4.** Esquema da validação cruzada, na qual os painéis 3K com (3K comercial), 3K cust (3K customizado), 7K com (7K comercial) e 7K cust (7K customizado) foram imputados para 50K.O grupo de validação foi alternado até que os cinco grupos (simbolizados por cada retângulo) tenham sido testados. Totalizando 20 cenários..... 33

## LISTAS DE TABELAS

### Capítulo I

<b>Tabela 1.</b> Números de marcadores (SNP) resultantes de painéis customizados e comerciais, após o controle de qualidade.....	35
<b>Tabela 2.</b> Acurácia da precisão dos genótipos imputados por indivíduo (ID) e por marcador (SNP).....	35

### Capítulo II

<b>Tabela 1.</b> Posição, em pares de bases, das regiões de maior efeito de Variância Aditiva Explicada para as características de comprimento de umbigo, pelagem e circunferência escrotal.....	48
<b>Tabela 2.</b> Números de acesso dos locos quantitativos posicionados em regiões cromossômicas que apresentaram associação com comprimento de umbigo e circunferência escrotal, e que possuem alguma relação, descrita na literatura, com esses traços.....	49
<b>Tabela 3.</b> Genes detectados pelo estudo de associação genômica ampla e suas funções biológicas, que foram identificadas pela análise de enriquecimento funcional.....	50

## RESUMO GERAL

A busca por plantéis cada vez mais produtivos, tem conduzido ao uso de diferentes ferramentas, dentre elas, a utilização de marcadores moleculares no melhoramento genético. No entanto, como o custo da técnica de genotipagem ainda é considerado elevado, uma alternativa é a utilização da técnica de imputação (predição de genótipos não interrogados na genotipagem). Sendo assim, no primeiro capítulo desta dissertação, analisou-se genótipos de animais Nelore com o objetivo de avaliar o potencial da imputação de painéis 3K e 7K, ambos customizados. Para a imputação foi utilizado o programa FImput. Os painéis foram produzidos a partir de marcadores que passaram pelo controle de qualidade, considerando a distância entre marcadores e os de maior frequência alélica. Os resultados demonstraram que a utilização de painéis customizados para imputação de marcadores, pode ser aplicada para alcançar valores de pelo menos 90% de acurácia. A customização do painel 3K e 7K aliada à técnica de imputação pode contribuir para melhorar o custo-benefício do uso de estudos genômicos no melhoramento animal. No segundo capítulo, objetivou-se realizar um estudo de associação genômica ampla (GWAS), por meio de dados imputados de animais da raça Canchim, visando identificar genes associados aos fenótipos de comprimento de umbigo, pelagem e circunferência escrotal. O GWAS foi realizado utilizando inferência *Bayesiana* por meio do programa GenSel. Foram detectadas 31 regiões em associação com comprimento de umbigo, uma região para pelagem e quatro regiões relacionadas à CE. Nessas regiões foram identificados locos quantitativos (QTLs) e genes relacionados com as características estudadas. A identificação de genes e QTLs, previamente descritos na literatura, reforçam as evidências de associação. Esse conhecimento poderá contribuir para a compreensão da arquitetura genética envolvida na expressão desses caracteres em bovinos Canchim.

**Palavras-chave:** Genômica; imputação; GWAS; Polimorfismos; Bovinos.

## **ABSTRACT**

The search for increasingly productive seedlings has led to the use of different tools, among them, the use of molecular markers in genetic improvement. However, since the cost of the genotyping technique is still considered high, an alternative is the use of the imputation technique (prediction of genotypes not interrogated in genotyping). Thus, in the first chapter of this dissertation, genotypes of Nellore animals were analyzed with the objective of evaluating the potential of imputation of 3K and 7K panels, both customized. For the imputation, the FImput program was used. The panels were produced from markers that passed the quality control, considering the distance between markers and those with higher allele frequency. The results demonstrated that the use of customized panels for imputation of markers can be applied to achieve values of at least 90% accuracy. The customization of the 3K and 7K panel together with the imputation technique can contribute to improve the cost-benefit of the use of genomic studies in animal breeding. In the second chapter, a broad genomic association study (GWAS) was carried out, using imputed data from Canchim animals, aiming to identify genes associated with navel length, coat and scrotal circumference phenotypes. GWAS was performed using Bayesian inference using the GenSel program. Thirty-one regions were found in association with umbilicus length, one coat region and four EC-related regions. In these regions, quantitative loci (QTLs) and genes related to the characteristics studied were identified. The identification of genes and QTLs, previously described in the literature, reinforce the evidence of association. This knowledge may contribute to the understanding of the genetic architecture involved in the expression of these characters in Canchim cattle.

**Keywords:** Genomic; Imputation; GWAS; Polymorphisms; Cattle.

## REVISÃO BIBLIOGRÁFICA

### Melhoramento genético

As raças bovinas podem ser divididas em dois sub-grupos, taurinos (*Bos taurus*) e zebuínos (*Bos indicus*), provenientes da Europa e Índia respectivamente. Esses animais apresentam 22.000 genes aproximadamente e diferem entre si em 237 posições genômicas (WILHAM, 2013).

Os taurinos foram trazidos para o Brasil em conjunto com outras espécies domesticadas por volta do ano de 1533. Os zebuínos foram introduzidos no Brasil em 1875, e se expandiram pelo país por apresentarem rusticidade, garantindo a sobrevivência desses animais mesmo em situações de estresse ambiental (da SILVA et al., 2012). Além disso apresentam boa adaptação a recursos alimentares a pasto e resistência à parasitas. Esses fatores, permitem explicar a numerosa população dos zebuínos que representam 80% dos rebanhos nacionais e sua importância para produção de carne e leite. No entanto, animais europeus se destacam em relação aos zebuínos quanto à aspectos de produtividade e reprodução (SANTANA et al., 2014).

Diante da necessidade de potencializar em bovinos características de adaptação, produtividade e reprodução, métodos de melhoramento genético para aumentar a frequência de genes associados à esses fenótipos tem sido empregados. O melhoramento genético clássico possui como objetivo elevar a cada ciclo de seleção a frequência de genótipos favoráveis com base em análises fenotípicas, obtendo-se assim um ganho genético de médio à longo prazo (ODA & MENCK, 1989).

Um dos métodos utilizados no melhoramento aplicado a bovinos considera ajustes em sistemas de produção, promovendo cruzamentos entre animais de diferentes raças, ou até mesmo entre *Bos taurus* e *Bos indicus*. Essa estratégia tem promovido vantagens para várias características de importância econômica, como por exemplo, aumento de peso e qualidade da carcaça. (PEROTTO et al., 2000; YOKOO et al., 2007).

Dentre as diferentes raças provenientes desses cruzamentos entre taurinos e zebuínos, destaca-se a raça Canchim. A raça europeia utilizada para este cruzamento foi a Charolesa e a zebuína inicialmente utilizada foi a Indubrasil. No entanto, outras raças zebuínas também foram incorporadas posteriormente, como a Nelore e Guzerá. Bovinos Canchim correspondem de maneira satisfatória às expectativas do produtor, pois possibilita melhor qualidade da carne, maior peso do rebanho, maior número de bezerras e rusticidade (adaptação ao clima e a parasitas) (BARBOSA, 2004).

Após diversas análises de esquemas para composição do “grau de sangue” do gado Canchim, foi definida a proporção de 5/8 Charolês e 3/8 zebuíno. Esse cruzamento favoreceu a presença de características desejáveis de produtividade e adaptação. Essa raça sintética tem se destacado dentre os pecuaristas, por apresentar-se como alternativa ideal para alavancar a produtividade dos plantéis sem o aumento da área de produção (ABCCAN, 2015).

Os bovinos Canchim tornaram-se relevantes para o produtor pela sua capacidade de disseminar características favoráveis ao rebanho brasileiro, possibilitados pelos chamados cruzamentos industriais, utilizando-se animais dessa raça. O cruzamento industrial é constituído por acasalamentos entre raças diferentes, buscando aumentar a eficiência produtiva dos animais, que serão levados à corte (MARCONDES et al., 2011; RIBEIRO et al., 2008).

Nesse contexto, a distribuição desses animais pelo país tem aumentado em média 8% ao ano, sendo que no ano de 2016 foram vendidos animais Canchim para quatro dos criatórios tradicionais do país, localizados em Mato grosso do Sul, Goiás, São Paulo e Minas Gerais (ABCCAN, 2016).

A Associação Brasileira de Criadores de Canchim (ABCCAN) destaca a boa receptividade dos criadores em relação a essa raça, como também a aptidão desses animais que podem se adaptar em qualquer região brasileira. No entanto, o principal obstáculo para que a disseminação desses animais ocorra de maneira mais intensa é questão logística, que dificulta a venda desses animais para locais distantes (ABCCAN, 2015).

Diante deste contexto, diversas técnicas de melhoramento têm sido aplicadas em busca da melhoria do desempenho desses animais, visando suprir a demanda crescente por carne. Uma das técnicas explorada para essa finalidade, é o uso de seleção assistida por marcadores moleculares, possibilitando ganhos genéticos em menor tempo, uma vez que propicia a identificação das bases genéticas envolvidas em determinados fenótipos de interesse (WILHAM, 2013).

### **Seleção assistida por marcadores moleculares**

O melhoramento genético utilizando marcadores moleculares tem proporcionado avanços na seleção de animais domésticos, seguindo o princípio de desequilíbrio de ligação (LD). O LD é caracterizado pela segregação em conjunto de genes e marcadores de maneira não aleatória, podendo ocorrer principalmente devido a proximidade física entre

ambos. Dessa maneira, por meio dos marcadores, busca-se identificar genes que possam ter influência em características de interesse econômico. Atualmente um dos marcadores mais utilizados para essa finalidade é o polimorfismo de nucleotídeo único (SNP, do inglês *Single Nucleotide Polymorphism*) (VIGNAL et al., 2002).

Os SNPs são também denominados de mutações pontuais, ocorrendo a troca de apenas um nucleotídeo, normalmente são encontrados apenas duas variantes em uma espécie, sendo chamados então de bi-alélicos. A mutação que caracteriza um SNP pode ser reconhecida como transição (troca de uma purina por purina ou pirimidina por pirimidina), ou por outra possibilidade menos frequente que são as trocas transversais (purina por pirimidina) (CAETANO et al., 2009).

O sequenciamento do genoma completo de diversas espécies permitiu reconhecer a abundância de marcadores do tipo SNP distribuídos por todo genoma. No entanto, para que esses sejam considerados SNPs devem estar presentes em pelo menos 1% da população testada. Os métodos de leitura de SNPs trouxeram novos benefícios e aplicações, com o desenvolvimento de painéis (também denominados “chips”) de genotipagem (BROOKES, 1999).

Os painéis desenvolvidos para a prospecção dos polimorfismos de base única, permitem a identificação de aproximadamente 777 mil marcadores quando realizados em chips de alta densidade (BovineHD BeadChip, HD). Além desses, existem também os menos densos, cujo custo é inferior, mas apresentam menor quantidade de marcadores identificados na análise (MARCHINI & HOWIE, 2010).

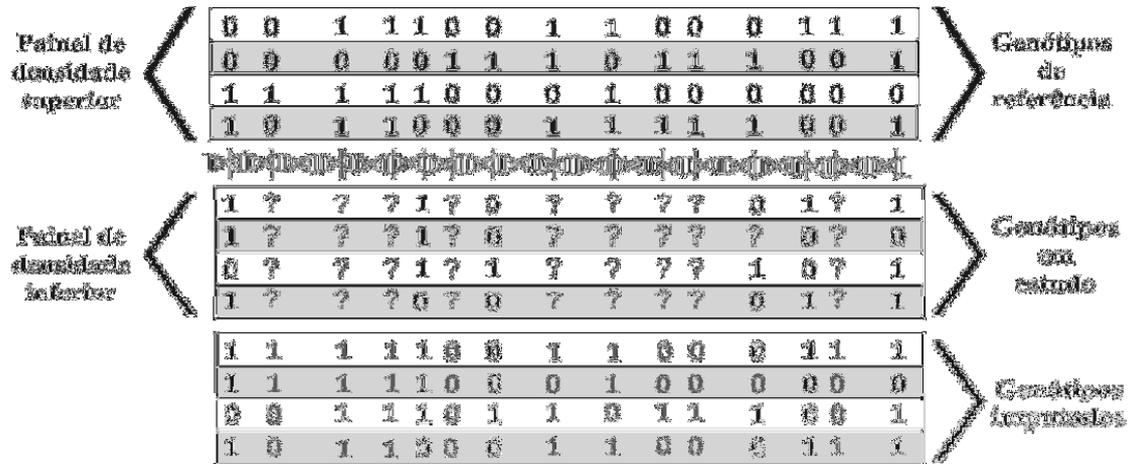
O elevado custo da técnica de genotipagem é uma limitação para estudos genômicos, no entanto maneiras alternativas para contornar esse problema têm sido sugeridas, como o desenvolvimento de metodologias de imputação (LI et al., 2009).

### **Imputação**

A ideia de imputação de genótipos foi descrita pela primeira vez por Burdick e colaboradores, que utilizaram o termo genotipagem *in silico*. Essa expressão foi utilizada para descrever que análises computacionais poderiam ser usadas para substituir procedimentos baseados em laboratório, na determinação de genótipos (LI et al., 2009).

A imputação é uma ferramenta utilizada para inferir genótipos não interrogados na genotipagem, tendo como base uma população de referência genotipada em painel de densidade superior (Figura 1).

### Cenário típico de imputação



**Figura 1.** Cenário típico de imputação. As barras na fita dupla DNA representam os marcadores (SNP). Os números (genótipos observados) e interrogação (? - genótipos não observados). Genótipos não obtidos no processo de genotipagem (?) no painel de densidade inferior são estimados os genótipos faltantes são determinados (Fonte: Nascimento, 2016).

As populações de referência, são assim denominadas por serem genotipadas em painéis mais densos. Dessa maneira possuem maior número de informação genética conhecida, possibilitando determinar os marcadores desconhecidos do grupo de validação (genótipo em estudo; população genotipada em menor densidade) (ZHANG et al., 2010; DASSONNEVILLE et al., 2012).

A precisão com que os genótipos são inferidos na imputação pode ser influenciada por vários fatores, dentre eles o tamanho das populações de referência e validação, a relação genética entre os animais, níveis de desequilíbrio de ligação e os métodos utilizados. Deste modo, cálculos para se obter acurácia de imputação são capazes de avaliar o quanto a técnica foi eficaz ao predizer os genótipos faltantes (JEWETT et al., 2012).

Devido à divergências no padrão de desequilíbrio de ligação de animais taurinos e zebuínos, supõe-se que painéis de alta densidade seriam capazes de realizar melhor cobertura do genoma zebuíno. Uma vez que os painéis tradicionalmente comercializados de baixa densidade possuem predominância de dados genéticos de animais taurinos. No entanto a utilização de painéis de alta densidade têm alto custo, inviabilizando muitos estudos com animais de origem indiana. Sendo assim, a técnica de imputação tem sido aliada à outras metodologias, como a customização de painéis de genotipagem

(ESPIGOLAN et al., 2013).

Os painéis customizados tendem a ser mais específicos em relação aos tradicionalmente utilizados. Um exemplo desses painéis é o *Z-chip seq* (~30 mil marcadores, 30K), o qual foi desenvolvido tendo como base informações genotípicas de animais Nelore, potencializando avaliação genética de animais de raças zebuínas. Assim como o *Z-chip*, outros painéis foram otimizados, com o intuito de possuírem informação genética predominante de animais de origem indiana, como o GeneSeek Genomic Profiler 80K (DASSONNEVILLE et al., 2012).

Ainda com o objetivo de tornar o uso de dados genômicos cada vez mais acessíveis, além dos painéis disponibilizados por empresas terceirizadas, existem metodologias de otimização, como a validação cruzada (do inglês *cross-validation*). Esse método baseia-se na divisão dos indivíduos em grupos e posterior análise em diversos ciclos sucessivos, de modo que cada grupo seja validado ao menos uma vez. Dessa maneira não se corre o risco de valores altos de acurácia, motivados por um dos grupos ser mais aparentados ao grupo de referência (REFAEILZADEH et al., 2008).

A customização de painéis de baixa densidade (custo reduzido) com marcadores para zebuínos, para posterior uso na imputação, é relevante uma vez que os painéis comercializados possuem informação genética predominante de animais europeus. Dentre os aspectos positivos da técnica, está a possibilidade de aumentar o número de animais com informações de marcadores mais densas, proporcionando avaliações genéticas de maior confiabilidade com menor custo (VENTURA et al., 2014).

Algumas das aplicações para a qual esses painéis de marcadores SNP têm sido explorados é o estudo de associação genômica ampla (GWAS, do inglês *Genome Wide Association Study*). Estes estudos tem como objetivo encontrar associações entre marcadores e genes e/ou locos que influenciam na expressão de fenótipos de interesse, utilizando ferramentas da bioinformática (ZHANG & DRUET, 2010).

### **GWAS – Estudos de Associação Genômica Ampla**

O GWAS parte do princípio de que marcadores abundantes no genoma, como os SNP, estejam em desequilíbrio de ligação com genes associados com características de interesse econômico. Para a realização do estudo, é necessário que haja tamanho amostral suficientemente grandes da população genotipada (que represente a variabilidade genotípica dos animais), para posterior análise estatística (CANTOR et al., 2010). Dessa maneira, a base dos estudos de associação genômica é a busca por combinações entre

informações genéticas e fenotípicas de indivíduos, com a finalidade de encontrar regiões que interferem em características que agreguem valor ao rebanho (HIRSCHHORN e DALY, 2005).

A principal importância dessa técnica de melhoramento, está na possibilidade de selecionar animais de acordo com informações de seu DNA, conduzindo a ganhos genéticos em menor período de tempo, quando comparado aos métodos tradicionais. Isso ocorre devido a vantagem de se analisar o animal logo após seu nascimento, uma vez que o mesmo possui em sua carga genética aqueles marcadores descritos em associação com fenótipos de interesse econômico, poderá permanecer no rebanho. Logo, os animais que não são portadores de genótipos favoráveis poderão ser descartados sem acarretarem prejuízos, causados pela criação de bovinos que não irão oferecer o retorno financeiro para o pecuarista (HIRSCHHORN e DALY, 2005).

Outro ponto positivo em relação à aplicação do GWAS no melhoramento genético, é a possibilidade de explorar características de baixa herdabilidade e difícil mensuração. No entanto, devido a necessidade de se manipular grande número de marcadores e de dados fenotípicos podem ocorrer nas análises erros do tipo I (falso positivo) e erros do tipo II (falso negativo). Para minimizar a possibilidade de que ocorram tais erros, são realizados procedimentos de controle de qualidade (CQ) nos quais, marcadores e amostras são testados de acordo com alguns parâmetros (CARMO, 2012).

## **Parâmetros do Controle de Qualidade dos marcadores**

### **Equilíbrio de Hardy-Weinberg**

O Equilíbrio de Hardy-Weinberg (EHW) determina que a frequência alélica permanece constante por muitas gerações nas situações em que: a população é grande, homogênea, com acasalamentos ao acaso e onde duas cópias do mesmo gene são herdadas independentemente. Esse parâmetro tem se consolidado no controle de qualidade, uma vez que, desvios no EHW podem indicar erros na determinação do perfil dos marcadores, endogamia, estratificação populacional ou seleção. Em populações em que os acasalamentos não ocorrem ao acaso, o EHW é utilizado para analisar se os resultados observados não se desviam totalmente do esperado para aquela população (CARMO, 2012).

## **Menor Frequência Alélica**

O parâmetro de menor frequência alélica (MAF) descreve a frequência do alelo menos comum na população. Alelos cuja frequência é inferior a 2% (valor geralmente utilizado para GWAS) dificultam o poder de associação entre fenótipo e genótipo. Este parâmetro é responsável pelo descarte de muitos SNPs das análises (GORLOV et al., 2008).

## **Parâmetros do Controle de Qualidade das amostras**

### **Call rate**

O *Call rate* refere-se a quantidade de genótipos determinados. Sendo uma medida para eliminar amostras com grande quantidade marcadores não interrogados. O limiar de exclusão geralmente utilizado, varia entre 95% a 99% (FARIA et al., 2014).

### **Heterozigose**

Valores elevados neste parâmetro podem ser indicativos de erros de manipulação, acarretando em contaminação do DNA, ou apenas indivíduos muito diferentes dos demais indivíduos da população. Para tanto, amostras a  $\pm 3$  desvios-padrão são excluídas visando contornar esse problema (FARIA et al., 2014).

### **Estratificação populacional**

A aplicação deste parâmetro objetiva avaliar a relação entre os indivíduos, tendo como base o parentesco genômico. Quando os indivíduos apresentam divergência nas populações ancestrais ocasiona no aumento de marcadores associados a fenótipos de maneira errônea (CARMO, 2012).

### **Método Estatístico**

Os métodos estatísticos para análise da relação causal entre o polimorfismo e as diferenças fenotípicas têm sido atualizados constantemente. Isso se deve à busca para quantificar e incorporar aos modelos, fontes de variações possíveis na população estudada, visando minimizar associações espúrias (KORTE et al., 2013). Os métodos estatísticos amplamente utilizados nos estudos de associação, dividem-se em inferência *Bayesiana* e

Frequentista (ou clássica), A principal diferença entre essas duas metodologias é que a estatística Bayesiana considera os parâmetros desconhecidos de um problema estatístico como variáveis aleatórias, enquanto a Frequentista os considera constantes (JÚNIOR et al., 2007).

Apesar das discussões a respeito da inferência *Bayesiana* terem se iniciado há séculos atrás, sua aplicação em genética quantitativa é relativamente recente, sendo introduzida por FERNANDO & GIANOLA (1986). Esses autores destacam sua potencialidade no melhoramento genético animal em geral. A estatística de *Bayes* foi proposta com o intuito de resolução de problemas em que o método frequentista apresentava limitações, sendo a principal delas as estimativas dos parâmetros, que levam em consideração muitas aproximações e forte suposição. As suposições citadas são em geral baseadas no teorema de limite central podendo levar a estimativas e previsões erradas, uma vez que é comum essas suposições serem violadas (JÚNIOR et al., 2005; BROOKS et al., 2011).

Em estatística *Bayesiana*, as inferências feitas sobre os parâmetros são oriundas da combinação entre informações *a priori* (componentes de variância) e os resultados das amostras, pelo teorema de *Bayes*. Essas informações *a priori* devem ser determinadas antes da análise dos dados, assumindo estimativas de distribuição das probabilidades para os parâmetros que não são conhecidos no modelo. A combinação das informações (*a priori* e amostral) fornecem resultados *a posteriori* dos parâmetros desconhecidos, que acarretará em informações sobre estes parâmetros após a observação dos dados (BROOKS et al., 2011).

Apesar da inferência *Bayesiana* ser considerada uma ferramenta poderosa para o melhoramento, alguns cálculos são intratáveis analiticamente, ou seja, não é possível sua resolução por métodos tradicionalmente utilizados. Sendo assim, é necessária a aplicação de modelos estocásticos (modelo matemático que utiliza distribuição de soluções associadas a uma probabilidade). Um dos métodos que pode ser adotado baseia-se em cadeias de Markov (MCMC, do inglês *Markov Chain Monte Carlo*) (JUNIOR et al., 2005).

MCMC é um método iterativo, pelo qual são feitas computacionalmente simulações em que cada algoritmo simulado depende da iteração anterior. Havendo iterações suficientemente grandes sua distribuição entrará em equilíbrio. Esse método visa solucionar problemas de integração e otimização computacional (GIANOLA & FERNANDO, 1986; GAMERMAN & LOPES, 2006).

Um dos programas utilizados para realização de análises Bayesianas utilizando cadeia de Markov é o GenSel (GARRICK & FERNANDO, 2009). O GenSel oferece opções de aplicação ao GWAS por meio dos métodos BayesB, BayesC $\pi$  e BayesC. Sendo a diferença entre elas a distribuição de probabilidades *a priori*.

## **Fenótipos**

As informações obtidas por meio dos resultados do GWAS podem melhorar a precisão ao selecionar animais com fenótipos de interesse. Para se realizar as análises de associação é importante a obtenção de informações fenotípicas da população estudada. Algumas características cuja forma de mensurar é facilitada, podem ser utilizadas as medidas originais, como por exemplo, medidas em centímetro. Para medidas de difícil mensuração são adotadas metodologias de escore, na qual são atribuídas notas ou pontuações de acordo com a forma ou tamanho da característica. Dentre as características para as quais se utilizam escore visual estão: comprimento de umbigo e pelagem (CANTOR et al., 2010; SANTANA et al., 2014).

## **Características Fenotípicas Avaliadas**

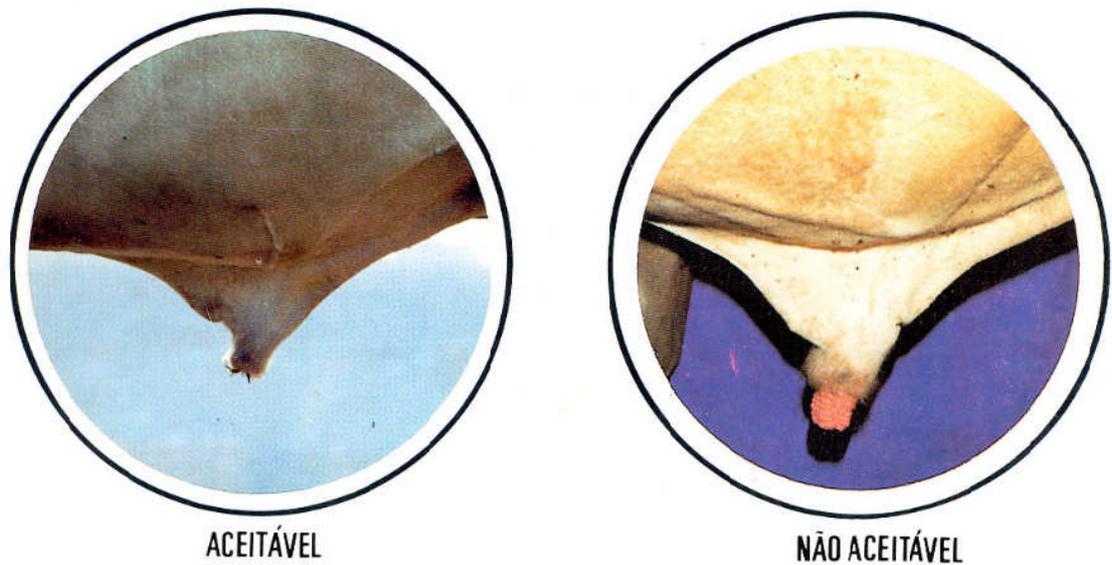
### **Comprimento de umbigo**

Animais que apresentam umbigo proeminente (distante da região ventral), estão mais propensos a lesões e patologias pelo contato direto com o solo, afetando a sanidade dos mesmos. Assim sendo, torna-se vantajoso a obtenção de animais com umbigo de menor tamanho (BIGNARDI et al., 2011). Nas fêmeas essa característica é analisada pela prega umbilical, e nos machos é considerado o conjunto prega umbilical e bainha.

A característica de umbigo pode ser considerada de importância econômica na criação de bovinos, pois problemas desse porte acarretam em gastos que variam de acordo com a gravidade das enfermidades. Perdas econômicas por óbito de animais, gastos com veterinário, medicamentos, retardo no crescimento e depreciação da carcaça são alguns dos prejuízos possíveis (ORTEGA et al., 2007; WINDEYER et al., 2014).

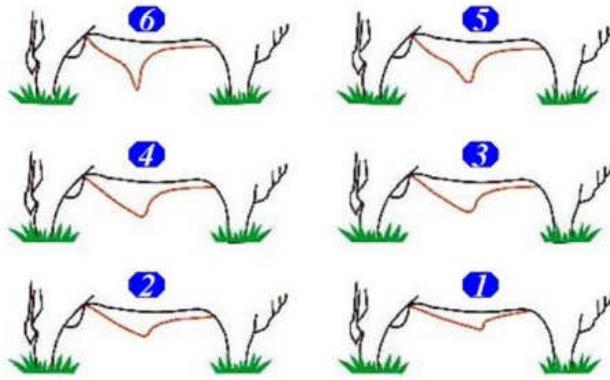
Devido a relação entre o umbigo e o prepúcio dos bovinos, o mesmo torna-se relevante para aspectos reprodutivos em touros, pois quando apresentam prolapso (ou relaxamento da bainha prepucial), sua mucosa fica exposta. Essa exposição permite que os animais estejam mais sujeitos à machucados na região (devido à sua sensibilidade), também podendo também acarretar em dificuldades na monta ou casos de infertilidade nos

machos. Sendo assim, é interessante que os touros apresentem prepúcio de menor tamanho, de maneira que a mucosa prepucial se posicione internamente à bainha até que haja exposição do pênis (Figura 2) (RABELO et al., 2012).



**Figura 2.** Mucosa prepucial interna à bainha (aceitável), mucosa prepucial exposta (não aceitável) (Fonte: Embrapa-São Carlos).

O fenótipo comprimento de umbigo é difícil de ser mensurado. Dessa forma, metodologias utilizando escores visuais são adotadas para viabilizar o estudo dessas características em programas de melhoramento (BARICHELLO et al., 2011). A avaliação visual dos animais visa pontuar as características de acordo com o tamanho ou forma do fenótipo estudado. Na Figura 3 está demonstrada a avaliação de umbigo, cujo escore pode variar de 1 a 6, sendo o de número 1 representado por umbigo colado a região ventral, os escores 2, 3 e 4 denominados funcionais, enquanto o 5 e 6 são considerados pendulosos. São preferíveis animais que apresentem umbigo próximo à região ventral (LEITE, 2009).



**Figura 3.** Método de avaliação por escore de comprimento de umbigo pelo sistema de avaliação de bovinos de 1 a 6 (Fonte: PPGZOO UFVJM, 2013).

### **Pelagem**

A característica pelagem (ou pelame) é classificada como aspecto racial bovino, definindo aqueles que mais se encaixam ao biotipo da raça. Além disso, a pelagem pode contribuir para adaptação dos animais (PORTO-NETO et al., 2014). Nesse sentido, o pelame foi descrito associado à infestação por parasitas. Estudos realizados em animais Canchim apontam a maior presença de carrapatos em animais de pelame mais escuro e comprido. Este fator pode estar relacionado a baixa imunidade desses animais em resposta ao calor (OLIVEIRA & ALENCAR, 1986).

A morfologia dos pelos bovino possuem importância no desempenho fisiológico dos animais diante dos fatores ambientais, características como comprimento do pelame, coloração, maior diâmetro, espessura da capa do pelame e densidade numérica, são aspectos que devem ser considerados, uma vez que o estresse térmico pode ser prejudicial para performance do rebanho (NICOLAU et al., 2004).

A coloração do pelame também apresenta associação positiva com a pigmentação da epiderme, sendo este, outro importante fator adaptativo, pois animais com pigmentação negra são menos permeáveis a radiação ultravioleta, tornando-se mais resistentes ao calor. Em ambientes caracterizados por elevada radiação, como regiões tropicais, é preferível criação de animais com pelame claro e epiderme negra (da SILVA et al., 2001; NICOLAU et al., 2004, Li et al., 2014).

A avaliação da característica de pelame bovino é realizada por escore visual. De acordo com o sumário de touros Canchim disponibilizado pela ABCCAN em parceria com

Embrapa-Genepplus, os escores para essa característica variam de 1 a 6, sendo mais indicados para região tropical animais de pelos curtos e claros (Embrapa- Genepplus, 2016).

### **Circunferência escrotal**

Outra característica importante para o melhoramento genético é a circunferência escrotal (CE), cujo fenótipo tem sido relacionado com desempenho reprodutivo dos touros. Essa relação é direta, ou seja, quanto maior a CE maior é a capacidade de realizar montas, podendo acarretar em maior número de bezerros no rebanho (FORDYCE et al., 2014; UTSONOMIYA et al., 2014)..

A alta herdabilidade encontrada para a característica CE (em torno de 0,50) favorece sua utilização em programas de melhoramento, por ser importante indicador de fertilidade e precocidade sexual (idade a puberdade) em reprodutores e novilhas. Sendo importante ressaltar sua correlação positiva com outros fenótipos de interesse, como por exemplo, peso e qualidade do sêmen. As medidas de circunferência escrotal são mensuradas em cm, sendo desejáveis touros de maiores medidas (PEREIRA et al., 2001; BOLIGON et al., 2007).

### **REFERÊNCIAS BIBLIOGRÁFICAS**

- 1.ABCCAN (2015) Canchim: cada vez mais forte. Disponível em: <http://www.abccan.com.br/canchim/index.php/noticias/materias-especiais/408-canchim-cada-vez-mais-forte.html>. Acesso em 26 setembro 2016
- 2.ABCCAN (2016) Raça Canchim. Disponível em: <http://www.abccan.com.br/canchim/index.php/a-raca.html>. Acesso em 5 junho 2016
3. Barbosa P (2004) A raça Canchim em cruzamentos para a produção de carne bovina. Embrapa Pecuária Sudeste Circular técnica
4. Barichello F, Alencar MMd, Torres Júnior RAdA, Silva LOCd (2011) Efeitos ambientais e genéticos sobre peso, perímetro escrotal e escores de avaliação visual à desmama em bovinos da raça Canchim. Revista Brasileira de Zootecnia:286-293
5. Bignardi A, El Faro L, Torres Júnior R, Cardoso V, Machado P, Albuquerque LGd (2011) Random regression models using different functions to model test-day milk yield of Brazilian Holstein cows. Genetics and Molecular Research:3565-3575
6. Boligon AA, Rorato PRN, Albuquerque Ld (2007) Correlações genéticas entre medidas de perímetro escrotal e características produtivas e reprodutivas de fêmeas da raça Nelore.

Revista Brasileira de Zootecnia 36 (3):565-571

7. Brookes AJ (1999) The essence of SNPs. *Gene* 234 (2):177-186

8. Brooks S, Gelman A, Jones G, Meng X-L (2011) *Handbook of Markov Chain Monte Carlo*. CRC press,

9. Caetano AR (2009) Marcadores SNP: conceitos básicos, aplicações no manejo e no melhoramento animal e perspectivas para o futuro. *Revista Brasileira de Zootecnia* 38 (8):64-71

10. Cantor RM, Lange K, Sinsheimer JS (2010) Prioritizing GWAS results: a review of statistical methods and recommendations for their application. *The American Journal of Human Genetics* 86 (1):6-22

11. Carmo AS (2012) Estudo de associação genômica e análise de enriquecimento genético da criotolerância espermática em bovinos da raça Nelore. Dissertação Universidade Estadual Paulista.

12. da Silva MC, Boaventura VM, Fioravanti MCS (2012) História do povoamento bovino no Brasil Central.

13. da Silva RG, La Scala Jr N, Pocay PLB (2001) Transmissão de Radiação Ultravioleta Através do Pelame e da Epiderme de Bovinos<sup>1</sup>. *Rev bras zootec* 30 (6):1939-1947

14. Dasonneville R, Fritz S, Ducrocq V, Boichard D (2012) Imputation performances of 3 lo-density marker panels in beef and dairy cattle. *American Dairy Science Association* 95:4136-4140.

15. Embrapa-São Carlos. Encurtando o umbigo. Disponível em: <http://ainfo.cnptia.embrapa.br/digital/bitstream/item/147064/1/FD01-Encurtando-o-umbigo.pdf>. Acessado em 20 de janeiro de 2017.

16. Embrapa-Geneplus. Sumário de touros Canchim, MA e Charolês. Disponível em: <http://geneplus.cnpqc.embrapa.br/sumarios/canchim/index.php?tp=download>. Acessado em 20 novembro 2016.

17. Espigolan R, Baldi F, Boligon A, Souza FRP, Gordo D, Tonussi RL, Cardoso DF, Oliveira HN, Tonhati H, Sargolzaei M, Schenkel FS, Carvalheiro R, Ferro JA, Albuquerque LG (2013) Study of whole genome linkage disequilibrium in Nelore cattle. *BMC genomics* 14:305.

18. Faria GMP, Pereira CS, Granato IS, Resende MDV, Silva FF, Ferreira KC, Rosse LN, Sansalone CP, Petrolí CD, Grattapaglia D (2014) Controle de qualidade de dados genotípicos para estudos genômicos em clones de eucaliptos. 7º Congresso brasileiro de melhoramento de plantas.

19. Fernando R, Gianola D (1986) Optimal properties of the conditional mean as a selection criterion. *Theoretical and applied genetics* 72 (6):822-825
20. Fordyce G, McGowan MR, Lisle A, Muller T, Allen J, Duff C, Holroyd RG, Corbet NJ, Burns BM (2014) Scrotal circumference of Australian beef bulls. *Theriogenology* 81 (6):805-812
21. Gamerman D, Lopes HF (2006) Markov chain Monte Carlo: stochastic simulation for Bayesian inference. CRC Press,
22. Garrick DJ, Taylor JF, Fernando RL (2009) Deregressing estimated breeding values and weighting information for genomic regression analyses. *Genetics Selection Evolution* 41 (1):1
23. Gorlov IP, Gorlova OY, Sunyaev SR, Spitz MR, Amos SI (2008) Shifting paradigm on association studies: value of rare single-nucleotide polymorphisms. *Am J Hum Genetics* 82(1):100-112.
24. Hartzell C, Putzier I, Arreola J (2005) Calcium-activated chloride channels. *Annu Rev Physiol* 67:719-758
25. Hirschhorn JN, Daly MJ (2005) Genome-wide association studies for common disease and complex traits. *Nature reviews genetics*. 6(2):95-108.
26. Jewett EM, Zawistowski M, Rosenberg NA, Zöllner S (2012) A coalescent model for genotype imputation. *Genetics* 191 (4):1239-1255
27. Wilham J (2013) A genômica bovina-origem e evolução de taurinos e zebuínos. *Veterinária e Zootecnia* 20 (2):217-237
28. Júnior JMC, de Assis GML, Euclides RF, Lopes PS (2005) Influência da informação a priori na avaliação genética animal utilizando dados simulados. *R Bras Zootec* 34 (6):1905-1913
29. Júnior JMC, TORRES RdA, LOPES PS (2007) Estimação de componentes de variância utilizando-se inferência Bayesiana e frequentista em dados simulados sob heterogeneidade de variâncias. *Revista Brasileira de Zootecnia: Brazilian Journal of Animal Science, Brasília* 36 (5):1539-1548
30. Korte A, Farlow A (2013) The advantages and limitations of trait analysis with GWAS: a review. *Plant methods* 9 (1):1.
31. Leite JF (2009) Método de avaliação visual EPMURAS na seleção e acasalamento dirigido de zebuínos. Universidade de São Paulo.
32. Li Y, Willer C, Sanna S, Abecasis G (2009) Genotype imputation. *Annual Rev Genomics and Human Genetics*. 10-:387-406.

33. Li M, Tiirikka T, Kantanen J (2014) A genome-wide scan study identifies a single nucleotide substitution in ASIP associated with white versus non-white coat-colour variation in sheep (*Ovis aries*). *Heredity* 112 (2):122-131
34. Marchini J, Howie B (2010) Genotype imputation for genome-wide association studies. *Nature Reviews Genetics* 11 (7):499-511
35. Marcondes MI, Valadares Filho S, Oliveira I, Paulino PVR, Valadares RFD, Detmann E (2011) Eficiência alimentar de bovinos puros e mestiços recebendo alto ou baixo nível de concentrado. *Revista Brasileira de Zootecnia* 40 (6):1313-1324
36. Nascimento W (2011) Estudos eletrofisiológico dos canais iônicos das células-tronco mesenquimais da geléia de Wharton do cordão umbilical humano. Universidade Federal de Pernambuco
37. Nascimento AV (2016) Genômica para prospecção de genes de interesse econômico em bovinos Nelore. Dissertação Universidade Federal da Grande Dourados
38. Nicolau C, Silva R, Mota L, Veríssimo C (2004) Características da pele e do pelame em bovinos da raça Caracu. *Arch Zootec* 53:25-34
39. Nunes P (2016) Fisiologia reprodutiva do macho bovino. Disponível em: [http://www.limousin.com.br/upload\\_artigos/FISIOLOGIA%20REPRODUTIVA%20DD%20MACHO%20BOVINO.pdf](http://www.limousin.com.br/upload_artigos/FISIOLOGIA%20REPRODUTIVA%20DD%20MACHO%20BOVINO.pdf). Acessado em 26 setembro 2016
40. Oda S, Menck A, Vencovsky R (1989) Problemas no melhoramento genético clássico do eucalipto em função da alta intensidade de seleção. IPEF: Instituto de pesquisas e Estudos Florestais, Piracicaba (41-42):8-17
41. Oliveira Gd, Alencar Md (1986) Resistencia de bovinos ao carrapato (*Boophilus microplus*).
42. Ortega J, Daft B, Assis R, Kinde H, Anthenill L, Odani J, Uzal FA (2007) Infection of internal umbilical remnant in foals by *Clostridium sordellii*. *Veterinary Pathology Online* 44 (3):269-275
43. Pereira E, Eler J, Costa F, Ferraz J (2001) Análise genética da idade ao primeiro parto e do perímetro escrotal em bovinos da raça Nelore. *Arq bras med vet zootec* 53 (1):116-121
44. Perotto D, Abrahao JdS, Moletta JL (2000) Características quantitativas de carcaça de bovinos Zebu e de cruzamentos *Bos taurus* x Zebu. *Revista Brasileira de Zootecnia* 29 (6):2019-2029
45. Porto-Neto LR, Kijas JW, Reverter A (2014) The extent of linkage disequilibrium in beef cattle breeds using high-density SNP genotypes. *Genetics Selection Evolution* 46 (1):1

46. PPGZOO UFVJM online (2013) Escores visuais corporais como auxílio à seleção em bovinos de corte. Boletim técnico. Disponível em: file:///C:/Users/NDREA~1/AppData/Local/Temp/Boletim%20T%C3%A9cnico%20PPPGZ%20v.1%20n%C2%B03%20F1%C3%A1vio%20Macedo.pdf. Acessado em 11 novembro 2016
47. Rabelo RE, Vulcani V, Cardoso LD, Campus Jataí G, Tavares H, Almeida P, Roncatto F (2012) Aspectos anatômicos e sua relação com as enfermidades do prepúcio e pênis no touro.
48. Refaeilzadeh P, Tang L, Liu H (2008) Cross-validation. Arizona State University.
49. Ribeiro ARB, Alencar M, Oliveira M (2008) Características do pelame de bovinos Nelore, Angus x Nelore e Senepol x Nelore. Reunião Anual da Sociedade Brasileira de Zootecnia 45
50. Santana MH, Utsunomiya YT, Neves HH, Gomes RC, Garcia JF, Fukumasu H, Silva SL, Junior GAO, Alexandre PA, Leme PR (2014) Genome-wide association analysis of feed intake and residual feed intake in Nelore cattle. BMC genetics 15 (1):1
51. Utsunomiya YT, Carmo AS, Neves HH, Carneiro R, Matos MC, Zavarez LB, Ito PK, O'Brien AMP, Sölkner J, Porto-Neto LR (2014) Genome-wide mapping of loci explaining variance in scrotal circumference in nelore cattle. PloS one 9 (2):e88561
52. Ventura R, Lu D, Schenkel F, Wang Z, Li C, Miller S (2014) Impact of reference population on accuracy of imputation from 6K to 50K single nucleotide polymorphism chips in purebred and crossbreed beef cattle. Journal of animal science 92 (4):1433-1444
53. Vignal A, Milan D, SanCristobal M, Eggen A (2002) A review on SNP and other types of molecular markers and their use in animal genetics. Genetics Selection Evolution 34 (3):1
54. Windeyer M, Leslie K, Godden S, Hodgins D, Lissemore K, LeBlanc S (2014) Factors associated with morbidity, mortality, and growth of dairy heifer calves up to 3 months of age. Preventive veterinary medicine 113 (2):231-240
55. Yokoo MJI, Albuquerque Ld, Lôbo RB, Sainz RD, Carneiro Júnior JM, Bezerra LAF, ARAUJO FdC (2007) Estimativas de parâmetros genéticos para altura do posterior, peso e circunferência escrotal em bovinos da raça Nelore. Revista Brasileira de Zootecnia 36 (6):1761-1768
56. Zhang Z, Druet T (2010) Marker imputation with low-density marker panels in Dutch Holstein cattle. Journal of Dairy Science 93 (11):5487-5494

## **HIPÓTESE**

1. Valores acima de 90% de acurácia são obtidos a partir da imputação de painéis customizados em animais Nelore.
2. Existe associação genômica ampla entre polimorfismos genéticos no DNA e os fenótipos de comprimento de umbigo, pelagem e circunferência escrotal em animais da raça Canchim.

## **OBJETIVO GERAL**

- Aplicação da técnica de imputação em animais Nelore e busca por associações entre características fenotípicas de interesse econômico e marcadores SNPs em bovinos da raça Canchim.

## **OBJETIVOS ESPECÍFICOS**

- Verificar 20 cenários de imputação em painéis customizados e comerciais
- Analisar a eficiência dos painéis customizados, quando submetidos à imputação para painel de maior densidade em animais Nelore;
- Prospectar genes associados com características de comprimento de umbigo, pelagem e circunferência escrotal em animais Canchim;
- Realizar o enriquecimento funcional, para a análise de genes com associação significativa com os fenótipos de interesse em animais Canchim.

## **CAPÍTULO I**

### **USO DE PAINÉIS COMERCIAIS E CUSTOMIZADOS DE POLIMORFISMOS DE NUCLEOTÍDEOS ÚNICOS PARA AVALIAÇÃO DA ACURÁCIA DE IMPUTAÇÃO**

Artigo em elaboração a ser submetido (Short Communication) :

**Genetics selection evolution**

**ISSN: 1297-9686**

**Fator de impacto: 2.895**

## Uso de painéis comerciais e customizados de marcadores SNP para avaliação da acurácia de imputação

### Resumo

O objetivo do trabalho foi avaliar acurácia de imputação de painéis Illumina® Bovine3K (3K) e Illumina® BovineLD (7K), ambos testados como comerciais (3k com e 7k com) e customizados (3k cust e 7k cust), utilizando dados genômicos de bovinos da raça Nelore. O grupo amostral foi composto por 522 animais genotipados em painel Illumina BovineSNP50 (50K) e 96 em Illumina BovineHD (HD). Animais genotipados em HD tiveram seus genótipos reduzidos para 50K e submetidos ao controle de qualidade. A customização dos painéis foi realizada utilizando os marcadores que passaram pelo controle de qualidade, considerando a distância entre marcadores e alelos de maior frequência. O método de validação cruzada utilizou 20 cenários de imputação. A acurácia da imputação foi estimada por porcentagem de alelos imputados corretamente (ACR), porcentagem de genótipos imputados corretamente (PERC) e correlação de *Pearson* ( $r$ ). A utilização de painéis de baixa densidade customizados submetidos à imputação foi eficiente, uma vez que os valores de acurácia para PERC e ACR obtidos ficaram em torno de 90%. Os painéis customizados (3k cust e 7k cust) apresentaram vantagem significativa em relação aos painéis comerciais (3k com e 7k com) em todos os cenários. A classificação dos painéis quanto a acurácia foi a mesma para PERC e ACR, tendo o 7k customizado como de melhor acurácia seguido pelo 3k customizado, 7k comercial e 3k comercial. Esse resultado indica que o uso de customização e imputação, possibilita diminuir custos de genotipagem e aumentar informações genômicas, proporcionando dessa forma maior confiabilidade e qualidade às análises subsequentes.

**Palavras-chave:** Genômica, *Bos indicus*, baixa densidade, desequilíbrio de ligação.

## **Introdução**

A disponibilidade dos painéis Illumina® Bovine3K (3K, ~3000 marcadores) e Illumina® BovineLD (7K, ~7000 marcadores) promoveram incentivos em estudos genômicos pela possibilidade de genotipar animais com menor custo. No entanto, sua baixa densidade dificulta a busca por associações entre genótipos e fenótipos de interesse. Outro impasse da utilização desses painéis está em sua aplicação em estudos com animais zebuínos, uma vez que as informações genômicas utilizadas como referência (para a composição dos painéis de genotipagem), foram predominantemente de animais de origem taurina [1,2].

Uma possibilidade de contornar as limitações do uso dos painéis comerciais é o desenvolvimento de painéis customizados de baixa densidade. Estes podem ser elaborados com base em genótipos de animais zebuínos, sendo mais específicos, de menor custo, e quando aliados às técnicas de imputação propiciam incremento considerável nos dados genômicos [3]. A imputação consiste na predição de genótipos que não foram interrogados na genotipagem. Para tanto, são inferidas informações genotípicas para indivíduos genotipados em menor densidade, tendo como referência animais genotipados em densidade superior [4].

Considerando a importância de raças zebuínas para a bovinocultura, torna-se necessário o uso de métodos alternativos que favoreçam o custo-benefício de estudos genômicos voltados ao melhoramento animal. Assim sendo, o objetivo do trabalho foi avaliar a acurácia de imputação de painéis 3K e 7K comerciais (3k com e 7k com) e customizados (3k cust e 7k cust) utilizando dados genômicos de bovinos da raça Nelore.

## **Material e Métodos**

### *Dados Amostrais*

Para a genotipagem das amostras, foram utilizados dois painéis de diferentes densidades, sendo eles: Illumina BovineSNP50 v2 (50K) e Illumina BovineHD (HD). Os dados foram lidos na plataforma *iScan* e, posteriormente, analisados usando o software R v.3.2.1 [5].

Para a realização das análises foram utilizados dados de 522 fêmeas Nelore (Fazenda Jatiúca - Água Clara/MS) genotipadas em painel 50K, e de 96 fêmeas Nelore (Agropecuária Jorge Ferreira - Juti/MS) genotipadas em HD. O grupo HD (96) foi reduzido para 50K, conforme o painel disponibilizado no site da empresa Illumina. Dessa

forma, totalizou-se dados de 618 animais em 50K que foram utilizados para realização dos testes de imputação.

### *Controle de Qualidade*

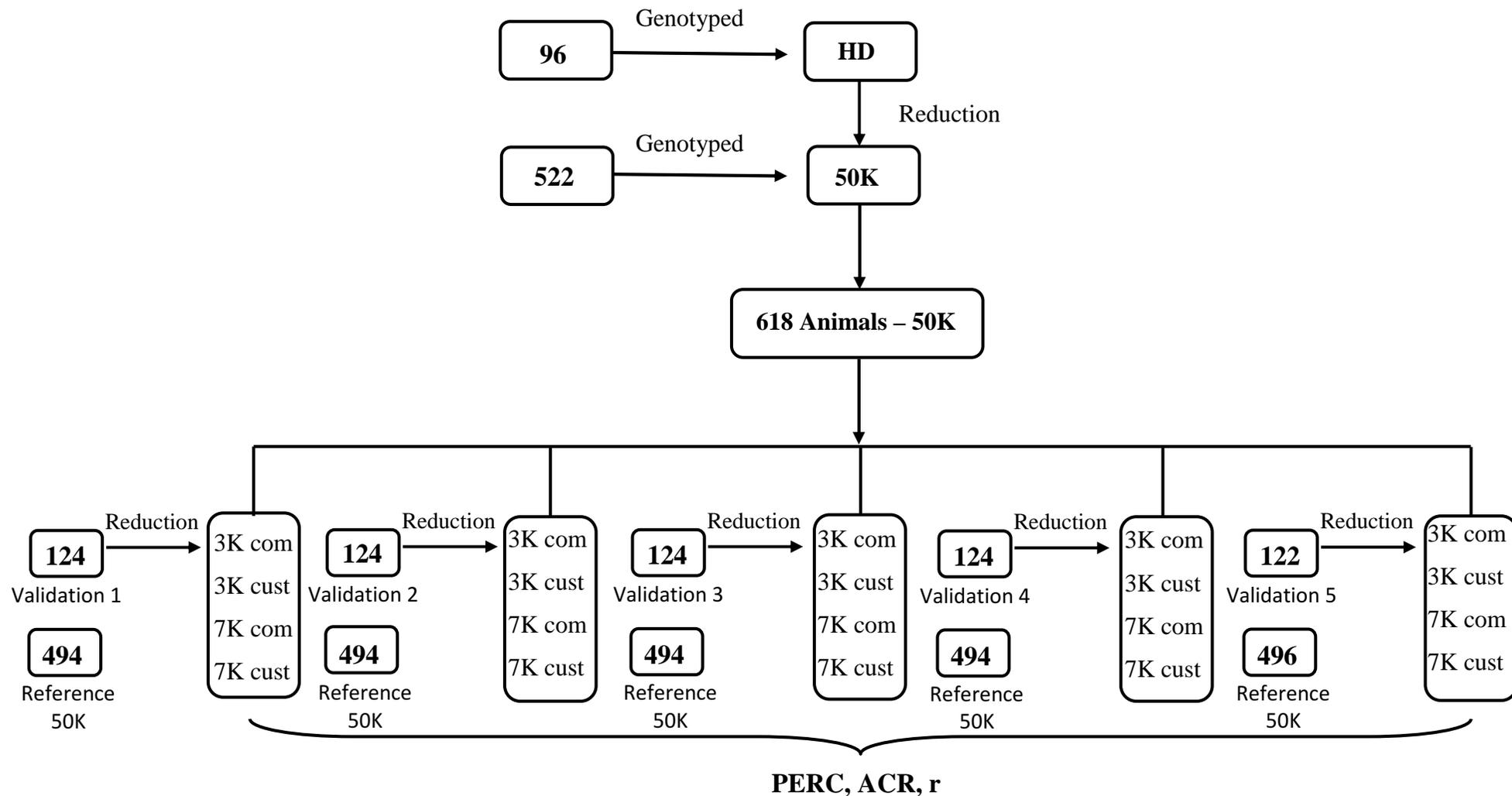
Os dados amostrais (618 animais 50K) foram submetidos ao controle de qualidade (QC, do inglês Quality Control), para a remoção de marcadores e amostras com possíveis erros de genotipagem ou duplicadas, sendo avaliadas por meio do pacote GenABEL v. 1.8-0 [6] integrado ao programa R v. 3.0.2. Os filtros de seleção utilizados foram: *Call rate* dos SNP (<95%), Equilíbrio de Hardy-Weinberg ( $P < 1e-05$ ), alelo menos frequente (MAF, do inglês *Minor Allele Frequency*) <2%, SNP de mesma coordenada genômica e remoção dos SNP não autossômicos. Os parâmetros de exclusão utilizados para indivíduos foram: *Call rate* (CR) inferiores a 90% e identidade por estado (IBS, do inglês *Identical By State*) maior que 95%.

Para imputação dos genótipos, que foi realizada pelo programa FImpute [7], painéis comerciais e customizados foram testados com densidade de 3K e 7K, para isso os dados em 50K foram mascarados. A customização foi realizada utilizando os marcadores que passaram pelo controle de qualidade, considerando a distância entre marcadores e os que tinham maior MAF dentro de janelas de cinco marcadores.

### *Cenários de Imputação*

A acurácia da imputação foi realizada de acordo com o método de validação cruzada, o qual realiza ciclos repetidos. Esse método baseia-se na divisão dos indivíduos em grupos e posterior análise em diversos ciclos sucessivos, de modo que cada grupo seja validado ao menos uma vez. Primeiramente, os dados foram agrupados, aleatorizados e divididos ao acaso em cinco grupos de igual tamanho, que em decorrência do número amostral, foram formados quatro grupos contendo 124 indivíduos e um grupo de 122.

Os conjuntos de referência e validação foram alternados de modo que cada um dos cinco grupos pudessem ser validados, totalizando 20 cenários de imputação. As populações de validação tiveram seu genótipo reduzido para 3K comercial (3K com), 3K customizados (3K cust), 7K comercial (7K com) e 7K customizados (7K cust). Os genótipos reduzidos para painéis comerciais, foram feitos a partir dos exemplares 3K e 7K disponibilizados no site da Illumina. Assim sendo, cada indivíduo teve seu genótipo mascarado para essas densidades ao menos uma vez (Figura 4).



**Figura 4** Esquema da validação cruzada, na qual os painéis 3K com (3K comercial), 3K cust (3K customizado), 7K com (7K comercial) e 7K cust (7K customizado) foram imputados para 50K. O grupo de validação foi alternado até que os cinco grupos (simbolizados por cada retângulo) tenham sido testados. Totalizando 20 cenários.

### Acurácia da Imputação

A acurácia da imputação foi estimada por porcentagem de alelos imputados corretamente (ACR, do inglês *Alleles Correctly*), porcentagem de genótipos imputados corretamente (PERC, sigla do inglês *Percentage of Correctly*) e correlação de *Pearson* ( $r$ ), comparando os genótipos imputados aos genótipos observados. Para o cálculo de ACR e PERC foram utilizadas as seguintes fórmulas, respectivamente:

$$ACR = \frac{\sum \left[ \frac{(p_i + 0.5 * pq_i)}{na_i} \right]}{N}$$

em que  $p$  é o número de genótipos com dois alelos imputados corretamente do indivíduo  $i$ ,  $pq$  número de genótipos com um alelo imputado corretamente do indivíduo  $i$ ,  $na$  número total de genótipos do indivíduo  $i$  e  $N$  número de indivíduos na população de validação.

$$PERC = \frac{\sum \left( \frac{gc_i}{ng_i} \right)}{N}$$

sendo  $gc$  número de genótipos imputados corretamente do indivíduo  $i$ ,  $ng$  número total de genótipos do indivíduo  $i$  e  $N$  número de indivíduos na população de validação.

Os cálculos foram realizados para indivíduo e para marcador. Os resultados para as diferentes densidades de painéis foram submetidos ao teste Kruskal-Wallis usando o pacote *agricolae* do R [8].

### Resultados e Discussão

Após o controle de qualidade, permaneceram 614 animais e 24066 marcadores que foram utilizados, posteriormente, para imputação.

Na Tabela 1, foram representados os números de marcadores utilizados nos painéis customizados e nos comerciais.

**Tabela 1** Números de marcadores (SNP) resultantes de painéis customizados e comerciais, após o controle de qualidade.

Painel	SNP
3K com	2025
3K cust	3012
7K com	4678
7K cust	7012
50K	24066

O painel 3K cust e 7K cust permaneceram com maior quantidade de marcadores após o controle de qualidade em relação ao comercial. Esse fato pode ser explicado, tanto pela eficiência da técnica de redução dos marcadores, como também pela predominância de genótipos de animais taurinos, utilizados como referência na confecção dos painéis de baixa densidade comercializados. A menor proporção de genótipos de referência de *Bos indicus* nos painéis comerciais 3K e 7K, pode ter determinado menor combinação de marcadores em comum durante o processo de redução dos genótipos [2].

Na Tabela 2 foram demonstradas as diferenças entre as acurácias dos painéis customizados e comerciais. O 7k customizado apresentou melhor acurácia seguido pelo 3k customizado, 7k comercial e 3k comercial.

**Tabela 2** Acurácia da precisão dos genótipos imputados por indivíduo (ID) e por marcador (SNP).

Painel	SNP			ID		
	PERC <sup>1</sup>	ACR <sup>2</sup>	r <sup>3</sup>	PERC <sup>1</sup>	ACR <sup>2</sup>	r <sup>3</sup>
3K com <sup>d</sup>	84,3 (0,106)	0,918 (0,057)	0,647 (0,181)	90,85 (0,039)	0,953 (0,022)	0,877 (0,036)
3K cust <sup>b</sup>	91,3 (0,06)	0,955 (0,032)	0,778 (0,143)	95,77 (0,029)	0,978 (0,016)	0,936 (0,024)
7K com <sup>c</sup>	90,8 (0,069)	0,953 (0,036)	0,8 (0,134)	84,32(0,030)	0,918 (0,016)	0,931 (0,025)
7K cust <sup>a</sup>	95,8 (0,036)	0,979 (0,019)	0,879 (0,109)	91,32 (0,018)	0,955 (0,01)	0,971 (0,014)

<sup>1</sup>Porcentagem de genótipos imputados corretamente; <sup>2</sup>Média de alelos imputados corretamente ;

<sup>3</sup>Correlação de Pearson; (Desvio padrão). Letras diferentes ao lado dos painéis indicam diferença estatística entre as acurácias.

Os painéis customizados apresentaram vantagem significativa em relação aos painéis comerciais em todos os casos. Os de acurácia PERC, ACR e r, apresentaram resultados semelhantes de modo que a classificação dos painéis foi a mesma para a maioria das avaliações. A média de alelos imputados corretamente apresentou valores elevados mesmo quando aplicados em painel de menor densidade (3K) (Tabela 2).

Estudos com animais zebuínos da raça Gir utilizando informações de pedigree, apresentaram valores semelhantes de ACR por marcador. O valor de 3K comercial obtido por Gir foi de 0,942 (desvio de 0,032), quanto ao de Nelore o ACR foi de 0,918 e 0,955 comercial e customizado, respectivamente [9].

Quanto ao 7K para Gir (0,97, desvio de 0,022) apresentou resultado superior ao de Nelore comercial e semelhante comparados ao customizado (0,953 e 0,979 comercial e customizado, respectivamente) [9]. Resultado similar de imputação para o cálculo PERC em painel 7K comercial foi obtido por Carvalheiro et al (2014) por SNP também com dados genômicos de Nelore (PERC de 90%) [10].

A correlação para indivíduo em Nelore alcançou vantagem dos painéis customizados 3K (0,936) e 7K (0,971), comparados aos de Boison et al (2015) em Gir. Sendo que este teve o valor de r para 3K de 0,92 (desvio-padrão de 0,053) e para 7K de 0,961 (desvio- padrão de 0,037).

Devido aos métodos de imputação considerarem o desequilíbrio de ligação entre os marcadores, espera-se que dados imputados, provenientes de animais taurinos tenham maiores valores de acurácia. Uma vez que os painéis comercializados possuem maior quantidade de informação genotípica de animais taurinos, favorecendo então a combinação de haplótipos compartilhados [11]. No entanto, neste trabalho foram encontrados valores de acurácia semelhantes aos de estudos realizados com bovinos das raças Holstein, Brown Swiss e Simental [13,14].

Para imputação de 3K para 50K, SARGOLZAEI et al (2011) realizando imputação de genótipos mascarados de animais europeus, encontraram valores de ACR correspondentes aos encontrados neste trabalho para painéis comerciais (0,95). Em relação ao 3K customizado, esses autores encontraram valores menores do que o 3K cust (0,97) aplicados em zebuínos. Esse resultado indica a potencialidade e importância de criação de protocolos customizados. É relevante ressaltar que os painéis comerciais de 3K e 7K apresentam maior porcentagem de informações genômicas oriundas de animais europeus.

Os resultados de  $r$  por indivíduo para as densidades 3K (0,936) e 7K (0,971) customizados também foram superiores quando comparados aos de DIMAURO et al (2013) utilizando o programa PSLR, com valores de 0,896 e 0,851 para 3K e para 7K, respectivamente. Utilizando o programa Beagle, o mesmo autor alcançou valores de 0,942 e 0,905, para 3K e 7K, respectivamente.

Os resultados indicaram que a acurácia oriunda da técnica de imputação pode variar de acordo com o painel utilizado (customizado ou comercial), a escolha dos métodos de imputação e cálculos de acurácia. Sendo assim, não há um protocolo que forneça precisão de imputação em todos os cenários [10].

Os resultados significativos para a acurácia neste estudo favorecem a aplicação das técnicas de customização e imputação em pesquisas relacionadas como a seleção genômica, uma vez que o valor genômico é inversamente proporcional à taxa de erro da imputação [15]. Sendo também possível a utilização destes painéis customizados em estudos de associação genômica (GWAS), levando a resultados mais confiáveis pelo maior número de marcadores [1].

## **Conclusão**

Os painéis customizados destacaram-se em relação aos comerciais com melhores resultados de acurácia. Esse fato indica a potencialidade do uso desses painéis para estudos genômicos, reduzindo custos e proporcionando dados com qualidade, aptos para aplicação em estudos posteriores.

## **Lista de abreviações**

**3K** – *Illumina Bovine3K*

**7K** – *Illumina BovineLD*

**50K** – *Illumina BovineSNP50*

**HD** – *Illumina BovineHD*

**3K cust** - *Illumina Bovine3K* customizado

**7K cust** - *Illumina BovineLD* customizado

**3K cust** - *Illumina Bovine3K* comercial

**7K cust** - *Illumina BovineLD* commercial

**ACR** – Alelos imputados corretamente, do inglês *alleles correctly*.

**PERC** – Porcentagem de genótipos imputados corretamente, do inglês *percentage of correctly*

**QC** – Controle de qualidade, do inglês Quality Control

**MAF** – Frequência do Alelo menor, do inglês do inglês *Minor allele frequency*

**IBS** – Identidade por estado, do inglês *Identical by state*

**SNP** - Polimorfismo de nucleotídeo único, do inglês *Single Nucleotide Polymorphism*

**r** – Correlação de *Pearson*

**GWAS** – Estudos de associação genômica ampla, do inglês *Genome Wide Association Study*.

### **Perspectivas Futuras**

O desenvolvimento de banco de dados nos quais pudessem ser disponibilizadas informações de dados genotipados em alta densidade, viabilizaria estudos na área de genômica. Assim como a disponibilização de painéis customizados, dessa forma seria possível a utilização de painéis de baixa densidade customizados (de menor custo), para que esses dados pudessem ser imputados.

### **Contribuição dos autores**

Andrea Romero contribuiu com a análise dos dados, discussão e a escrita do manuscrito inicial. André Nascimento, Yuri Utsonomyia e Adam Utsonomyia contribuíram com a manipulação e análise dos dados e o gerenciamento das ferramentas utilizadas para a idealização do trabalho. Alexéia Grisolia forneceu os dados manipulados e orientou as atividades. Fernando Garcia forneceu suporte e contribuiu com o planejamento do projeto idealizado.

### **Agradecimentos**

Agradecemos à Universidade Federal da Grande Dourados (UFGD) e à equipe do Prof. Dr Fernando Garcia da Faculdade de Medicina Veterinária, UNESP, Araçatuba-SP.

### **Conflito de interesses**

Os autores declaram que não possuem interesses concorrentes.

### **Apoio financeiro**

Esse trabalho teve o suporte da Fundação de Apoio ao Desenvolvimento do Ensino, Ciência e Tecnologia do Estado de Mato Grosso do Sul (FUNDECT) e Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

## Referências

1. Bush WS, Moore JH. Genome-wide association studies. *PLoS Comput Biol.* 2012;8 (12):e1002822
2. Weigel K, Van Tassell C, O'Connell J, VanRaden P, Wiggans G. Prediction of unobserved single nucleotide polymorphism genotypes of Jersey cattle using reference panels and population-based imputation algorithms. *Journal of Dairy Science.* 2010;93 (5):2229-2238
3. Dassonneville R, Fritz S, Ducrocq V, Boichard D. Short communication: Imputation performances of 3 low-density marker panels in beef and dairy cattle. *Journal of dairy science.* 2012;95 (7):4136-4140
4. Druet T, Schrooten C, De Roos A. Imputation of genotypes from different single nucleotide polymorphism panels in dairy cattle. *Journal of Dairy Science.* 2010;93 (11):5443-5454
5. R Development Core Team. R Foundation for Statistical Computing, Vienna, Austria. 2008
6. Aulchenko YS, Ripke S, Isaacs A, Van Duijn CM. GenABEL: an R library for genome-wide association analysis. *Bioinformatics.* 2007;23 (10):1294-1296
7. Sargolzaei M, Chesnais JP, Schenkel FS. A new approach for efficient genotype imputation using information from relatives. *BMC genomics.* 2014;15 (1):1
8. de Mendiburu F. Tutorial de agricolae. 2016;(Versión 1.2-4).
9. Boison S, Santos D, Utsunomiya A, Carvalheiro R, Neves H, O'Brien AP, Garcia J, Sölkner J, da Silva M. Strategies for single nucleotide polymorphism genotyping to enhance genotype imputation in Gyr (*Bos indicus*) dairy cattle: Comparison of commercially available single nucleotide polymorphism chips. *Journal of Dairy Science.* 2015;98 (7):1-21
10. Carvalheiro R, Boison SA, Neves HH, Sargolzaei M, Schenkel FS, Utsunomiya YT, O'Brien AMP, Sölkner J, McEwan JC, Van Tassell CP. Accuracy of genotype imputation in Nelore cattle. *Genetics Selection Evolution.* 2014;46 (1):1
11. Porto-Neto LR, Kijas JW, Reverter A. The extent of linkage disequilibrium in beef cattle breeds using high-density SNP genotypes. *Genetics Selection Evolution.* 2014;46 (1):1
13. Dimauro C, Cellesi M, Gaspa G, Ajmone-Marsan P, Steri R, Marras G, Macciotta NP. Use of partial least squares regression to impute SNP genotypes in Italian Cattle breeds. *Genetics Selection Evolution.* 2013;45 (1):1

14. Sargolzaei M, Schenkel F, Chesnais J Accuracy of imputed 50k genotypes from 3k and 6k chips using FImpute version 2. In: Dairy Cattle Breeding and Genetics Committee Meeting, 2011. pp 1-9
15. Khatkar MS, Moser G, Hayes BJ, Raadsma HW. Strategies and utility of imputed SNP genotypes for genomic analysis in dairy cattle. BMC genomics. 2012;13 (1):1

## **CAPÍTULO II**

### **Prospecção de genes associados com características de umbigo, pelagem e circunferência escrotal em bovinos Canchim**

Artigo em elaboração a ser submetido

(Research paper)

**Livestock Science**

**ISSN: 1871-1413**

**Fator de impacto: 1.377**

## **Prospecção de genes associados com características de comprimento de umbigo, pelagem e circunferência escrotal em bovinos Canchim**

### **Resumo**

O objetivo desse trabalho foi identificar regiões genômicas e genes em associação com as características de comprimento de umbigo, pelagem e circunferência escrotal (CE) em bovinos da raça Canchim. Foram utilizados dados imputados de marcadores do tipo polimorfismos de nucleotídeo único, com acurácia superior a 95%. As análises de associação genômica ampla foram conduzidas pelo método Bayes B implementado no programa GenSel, considerando como significativas às regiões que apresentaram Variância Genética Aditiva Explicada superior a 0,19%. O GWAS indicou 31 regiões em associação com comprimento de umbigo, uma região para pelagem e quatro regiões relacionadas à CE. Nessas regiões foram identificados 4 locos quantitativos (QTLs) relacionados com comprimento de umbigo e 5 com CE, quanto a pelagem os QTLs detectados não apresentaram relação, previamente descrita, com essa característica. Dentre os genes candidatos nas regiões associadas, 7 genes foram citados em trabalhos anteriores, com funções biológicas relacionáveis a característica de umbigo, um para pelagem e três para CE. Os genes *TMEM176A* e *TMEM176AB*, candidatos à associação com a característica umbigo, são ortólogos em humanos e constituem o cluster *Androgen Induced 1*. A identificação de genes e QTLs, previamente descritos na literatura, com função relacionada com as características estudadas reforçam as evidências de que as regiões encontradas estão associadas com as características de comprimento de umbigo, pelagem e CE. Esse conhecimento poderá contribuir para a compreensão da arquitetura genética envolvida na expressão desses caracteres em bovinos Canchim.

**Palavras-chave:** Inferência Bayesiana. GWAS. Polimorfismos. SNP.

*Financiamento:* Esse trabalho foi financiado pela Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA).

## **1. Introdução**

Rebanhos destinados à produção de carne que utilizam touros Canchim (5/8 Charolês e 3/8 zebuínos) em sistemas de acasalamento podem obter ganhos que chegam a 2,7@ no peso da carcaça ao abate em comparação com demais bovinos de corte (Tupy et al., 2006). Essa raça também tem se destacado com sua qualidade de carne e por possui boa adaptação às condições de ambiente tropical (Buzanskas et al., 2014; Mokry et al., 2013). Pesquisas de genética quantitativa já foram conduzidas nessa raça para avaliar parâmetros genéticos de características relacionadas ao desempenho reprodutivo, escore visual de umbigo e pelagem. No entanto, as regiões específicas ou genes responsáveis pela variação genética dessas características ainda devem ser elucidadas (Barichello et al., 2010; Bignardi et al., 2011; Buzanskas et al., 2014).

Estudos de associação genômica ampla (GWAS) têm sido utilizados em bovinos de corte para avaliar marcadores do tipo polimorfismos de nucleotídeo único (SNP), envolvidos na manifestação de características quantitativas. As informações obtidas a partir desses marcadores podem explicar parte da variação nos fenótipos de interesse econômico (Cantor et al., 2010; Santana et al., 2014).

Dentre os vários fenótipos que são de interesse para o melhoramento do rebanho, estão àqueles voltados à adaptação e funcionalidade, como por exemplo, pelagem e comprimento de umbigo. Animais com umbigo que ultrapassem a linha do jarrete estão mais susceptíveis a prolapso de prepúcio e patologias por terem maior contato com o pasto, portanto, é conveniente eliminá-los do plantel produtivo. A característica de pelame está ligada a aspectos raciais e adaptativos, sendo preferíveis, para regiões tropicais, bovinos com pelos curtos e claros (Ortega et al., 2007; Windeyer et al., 2014).

Outra característica que afeta economicamente a criação de bovinos de corte é a avaliação da circunferência escrotal (CE), que está diretamente relacionada com o desempenho reprodutivo em touros e precocidade sexual nas filhas desses touros. Essa característica é um importante indicador de fertilidade, por apresentar correlação genética com precocidade sexual em reprodutores e novilhas (Fordyce et al., 2014; Utsunomiya et al., 2014).

Neste contexto e visando contribuir com o melhoramento genético da raça Canchim, o objetivo deste trabalho foi identificar genes e regiões cromossômicas envolvidas com as características comprimento de umbigo, pelagem e CE por meio de métodos Bayesianos.

## **1. Material e Métodos**

### **2.1. Dados de valores genéticos preditos desregressados**

Para avaliação das características foram utilizados valores genéticos preditos desregressados (dEBV) de 1106 animais para as características comprimento de umbigo, pelagem e circunferência escrotal, por meio da metodologia proposta por Garrick et al. (2009). Essas informações foram obtidas do banco de dados do Programa Geneplus/Embrapa de Melhoramento Genético da raça Canchim.

### **1.2. Dados genotípicos**

Os genótipos dos animais estudados foram provenientes do banco de dados da Embrapa Beef Cattle (EBC) e Embrapa Southeast Livestock (ESL). Quanto aos dados provenientes da EBC, o primeiro conjunto de dados contendo 462 indivíduos foram genotipados por meio do Illumina *BovineSNP50 v2 BeadChip*® (50K). O segundo conjunto de dados foi composto por 169 indivíduos genotipados por meio do *GeneSeek Genomic Profiler HD*® (80k). O terceiro conjunto de dados foi composto por 13

indivíduos genotipados por meio do *BovineHD BeadChip*® (HD). Desse total de 644 animais genotipados, 548 eram da raça Canchim e 96 do grupo genético MA (progênie de touro Charolês e vaca ½ Canchim + ½ Zebu e apresentando em média 65,6% Charolês e 34,4% Zebu).

Todos os genótipos do banco de dados da EBC pertencem a animais que participaram das Provas Canchim de Avaliação de Desempenho (PCADs), avaliados entre os anos de 2011 à 2014. As PCADs são promovidas pela Associação Brasileira dos Criadores de Canchim (ABCCAN) e Programa Geneplus/Embrapa.

Referente ao banco de dados genômicos da ESL foram utilizadas informações de 57 animais (filhos de 49 touros) que possuíam DEP catalogada. Estes marcadores foram genotipados em painel HD, sendo 30 fêmeas (21 Canchim e 9 MA) e 27 machos (11 Canchim e 16 MA).

### 1.3. Controle de Qualidade e Imputação

Foram considerados somente SNPs com coordenadas genômicas conhecidas e contidos nos cromossomos autossômicos e, nesta etapa, cada conjunto de dados foi analisado separadamente. Os seguintes critérios e limiares de exclusão foram utilizados: genótipos com *Genotype Calling Score* inferior a 50%; amostras e SNPs com *Call Rate* inferior a 95%; frequência alélica mínima (MAF) de 5%; amostras que apresentaram heterozigose > 75%; valor de P do teste qui-quadrado para equilíbrio de Hardy-Weinberg < 0.000001; para SNPs situados em mesma coordenada genômica excluiu-se o de menor MAF; para pares de SNPs com correlação superior a 0,98, excluiu-se o de menor MAF. Essa etapa foi realizada por meio de *scripts* no *software* estatístico R v.3.1.3 (R Development Core Team, 2013), utilizando o pacote *snpStats* v1.16.0 (Clayton, 2012).

A imputação foi realizada por meio do *software* FImpute v.2.2. (Sargolzaei et al., 2014), de acordo com os procedimentos descritos por Santiago et al. (2017). Após a imputação, os dados genômicos foram submetidos a um novo controle de qualidade, em que permaneceram para os estudos posteriores apenas aqueles marcadores com acurácia maior do que 95%, MAF maior que 5% e  $P > 0.000001$  após o teste de qui-quadrado para o equilíbrio de Hardy Weinberg. SNPs com valor inferior de MAF entre pares de SNP altamente correlacionados ( $> 0,98$ ) foram excluídos.

#### 1.4. Estudo de Associação genômica ampla (GWAS)

O GWAS foi conduzido por meio do sistema operacional Linux, utilizando o método *Bayes B* contido no programa GenSel v.1.14 (Fernando and Garrick, 2008), utilizando simulação estocástica Monte Carlo via cadeias de Markov para estimar o efeito dos marcadores.

Os dados de genótipos e dEBV foram simultaneamente analisados, sendo utilizados apenas animais que apresentassem ambas informações. Dessa maneira, após o controle de qualidade totalizaram 697 animais (com informações de genótipos e dEBV), usados para análises subsequentes.

As cadeias de Markov incluíram 50.000 iterações, cujas primeiras 10.000 foram descartadas no *burn-in*. Estimativas dos efeitos de marcadores foram amostradas a cada quatro iterações. A inferência de associação nesse modelo de regressão múltipla foi realizada com base em janelas de 1 Megabase (Mb). A distribuição de  $\pi$  utilizada *a priori* em *Bayes B* foi obtida por meio da metodologia de *BayesC $\pi$* , utilizando o valor *a priori* de 0,5 (Habier et al., 2011).

Janelas que contêm vários marcadores capturam uma grande parte da variabilidade dentro das regiões cromossômicas, portanto, a associação entre regiões

genômicas e traços fenotípicos foi testada através da avaliação de 2.521 janelas não sobrepostas de 1Mb (1 milhão de pares de bases) (Saatchi et al., 2012). A proporção esperada de variância genética aditiva para uma janela específica foi calculada de acordo com Onteru et al. (2013), em que  $100\% / 2.521$  resultaram em 0,039%. Assim, as janelas que explicam pelo menos 0,19% da variância genética, que é cinco vezes maior do que o esperado ( $0,039\% * 5 = 0,19\%$ ), foram consideradas regiões associadas à QTLs. O efeito dos marcadores foi padronizado de acordo com Sollero et al. (2014), em que  $(a_i)^2/\alpha^2_{ai}$ .

### 1.5. Enriquecimento Funcional

As regiões significativas detectadas pelo GWAS foram exploradas por meio do banco de dados *cattleQTLdb* (Hu et al., 2006), para reconhecer locos de caracteres quantitativos (QTLs) que foram descritos em bovinos. Essas regiões candidatas a associação foram também submetidas a buscas no banco de dados BioMart (Ensembl Release 89) (Kinsella et al., 2011) utilizando a anotação do genoma bovino UMD3.1, para a procura dos genes nessas regiões identificadas. Para investigar associações já descritas, entre as características estudadas e os genes identificados pelo GWAS, foram realizadas buscas na literatura disponível.

Os genes detectados foram submetidos à análise funcional, utilizando o programa DAVID Bioinformatics Resources 6.8 (Dennis et al., 2003). Por meio do DAVID, buscou-se interações significativas ( $P < 0,001$ ) entre os genes e *clusters* relacionados com as características de umbigo, pelagem e CE, para essa busca utilizou-se informações de genes ortólogos em humanos, devido sua melhor anotação do genoma.

## 2. Resultados

O GWAS identificou regiões significativamente associadas com as características estudadas ( $VGAE > 0.19$ ), sendo 31 para comprimento de umbigo, uma região para pelagem e quatro regiões relacionadas à CE. Devido ao elevado número de regiões identificadas para a característica comprimento de umbigo, somente foram utilizadas na busca funcional as 10 regiões com maior significância (Tabela 1).

**Tabela 1.** Posição de início e fim, em pares de bases, das regiões de maior efeito de Variância Aditiva Explicada (VAE) para as características de comprimento de umbigo, pelagem e circunferência escrotal.

BTA <sup>1</sup>	Característica	Posição (pb) <sup>2</sup>	Número de SNPs	VAE <sup>3</sup> (%)
4	Umbigo	114054220 - 114991337	147	1.1
14	Umbigo	34002686 - 34989681	145	1.0
5	Umbigo	61006403 - 61985977	160	0.9
29	Umbigo	29000227 - 29997975	152	0.7
22	Umbigo	22012223 - 22999992	194	0.6
15	Umbigo	19003159 - 19982187	143	0.5
1	Umbigo	141006408 - 141995096	139	0.5
5	Umbigo	49004461 - 49987641	111	0.4
9	Umbigo	89003986 - 89993018	104	0.4
3	Umbigo	17052592 - 17933808	200	0.4
3	Pelagem	36000665 - 36995924	108	1.9
8	CE	106001959 - 106996237	97	0.3
21	CE	35026794 - 35999106	210	0.2
12	CE	28013232 - 28999952	137	0.2
1	CE	71001536 - 71997868	181	0.2

<sup>1</sup>Cromossomo bovino; <sup>2</sup>pb: posição onde a janela começa e termina em pares de bases;

<sup>3</sup>Variância aditiva explicada.

A busca por QTLs no banco de dados *CattleQTLdb* resultou na identificação de 68 QTLs nas janelas referentes à umbigo, 4 para pelagem e 39 para CE. Desses valores totais de QTLs, aqueles que foram previamente descritos relacionados com as características abordadas estão listados na tabela 2. As regiões genômica mais

promissoras associadas ao comprimento de umbigo foram localizadas no cromossomo 5. Apesar de não haverem locos específicos para umbigo, foram identificados QTLs para características previamente relacionadas. Para característica de CE os QTLs identificados são relacionados à reprodução em machos e fêmeas. Nas janelas em associação com pelagem, foram detectados QTLs, porém estes não foram relacionáveis com a característica de pelame.

**Tabela 2.** Números de acesso dos locos quantitativos posicionados em regiões cromossômicas que apresentaram associação com comprimento de umbigo e circunferência escrotal, e que possuem alguma relação, descrita na literatura, com esses traços.

<b>Característica</b>	<b>QTLs<sup>1</sup></b>	<b>BTA<sup>2</sup></b>	<b>Número de acesso<sup>3</sup></b>
<b>Umbigo</b>	Coloração da pelagem	5	37341
	Taxa de Fertilidade	5	62357
<b>Circunferência Escrotal</b>	Taxa de prenhes das filhas	8	43818
	Facilidade no parto	8	43825
	Concepção por inseminação	12	15276
	Quantidade espermática	1	65823

<sup>1</sup>Loci de caracteres quantitativos. <sup>2</sup>Cromossomo bovino, <sup>3</sup>Número de acesso no Cattle QTLdb

O banco de dados *Ensembl*, consultado para explorar as regiões significativas, possibilitou identificar 88 genes referentes ao comprimento do umbigo, dois para a característica pelagem e 37 para CE (Table S2). Esses genes candidatos foram submetidos à busca na literatura disponível, em que 7 genes foram citados em trabalhos anteriores, com funções biológicas relacionáveis a característica de umbigo, um para pelagem e três para CE. Esses genes estão demonstrados na tabela 3.

**Tabela 3.** Genes detectados pelo GWAS e suas funções biológicas identificadas pela análise de enriquecimento funcional.

Gene	Fenótipo	Cromossomo	Função	Nome do Gene
<b>IGSF5</b>	Umbigo	1	Reprodução	immunoglobulin superfamily member 5
<b>NOS3</b>	Umbigo	4	Prepúcio	nitric oxide synthase 3
<b>TMEM176A</b>	Umbigo	4	Pele	transmembrane protein 176A
<b>TMEM176B</b>	Umbigo	4	Pele	transmembrane protein 176B
<b>WIF1</b>	Umbigo	5	Pelagem	WNT inhibitory factor 1
<b>PKIB</b>	Umbigo	9	Adaptação	Proteína kinase
<b>DDX25</b>	Umbigo	29	Reprodução	DEAD-box helicase 25
<b>NTNG1</b>	Pelagem	3	Pelagem	Netrin G1
<b>BRCA2</b>	CE	12	Reprodução	DNA repair associated
<b>LOXL1</b>	CE	21	Reprodução	lysyl oxidase like 1
<b>GZMB</b>	CE	21	Reprodução	granzyme B

Os genes *TMEM176A* e *TMEM176AB*, candidatos à associação com a característica umbigo, são ortólogos em humanos e constituem o cluster *Androgen Induced 1* (AIG1, Id: GO:0005525). Em relação à pelagem e CE não foram identificadas relações entre os genes candidatos e clusters.

### 3. Discussão

Os QTLs presentes em regiões associadas ao fenótipo umbigo, que são relativos à reprodução, corroboram com estudos anteriores que destacam a importância do comprimento do umbigo para aspectos reprodutivos em touros. Isso ocorre devido à fatores hereditários que predispõem animais de umbigos alongados e pendulosos a terem prolapso prepucial (Koury Filho et al., 2003).

Mediante a relação previamente estabelecida entre comprimento do umbigo e prepúcio, o gene *NOS3*, candidato a associação com a característica de umbigo, foi identificado em células endoteliais microvasculares dérmicas de prepúcio neonatal. Essas células possuem atividades em processos fisiológicos e patológicos na pele (Cha et al., 2005; McGowan et al., 2002).

Ainda com relação às regiões em associação com umbigo, os genes *IGSF5* e *DDX25* também apresentaram relação com aspectos reprodutivos. O gene *IGSF5* foi descrito como associado à infertilidade em estudos de expressão utilizando ratos. Este gene é expresso no mesonefro, necessário para a formação dos túbulos seminíferos. Estes túbulos situam-se no testículo, no qual ocorre a produção dos espermatozoides (Gitton et al., 2002). O gene *DDX25* está presente em células testiculares de roedores e sua expressão apresentou-se como essencial para a fertilidade de camundongos (Tsai-Morris et al., 2007).

Umbigo e pelagem são importantes características físicas para a adaptação de bovinos. Sendo assim, regiões genômicas em comum que estejam afetando a expressão desses fenótipos podem ser benéficos para o melhoramento de ambos concomitantemente (Corbet et al., 2007). Dessa maneira, no BTA 5 do GWAS para característica de comprimento de umbigo, foi identificado QTL relacionado com o fenótipo de pelagem, Porto-Neto et al. (2014) estudando animais da raça Brahman, também identificou a relação entre esses dois traços, neste mesmo cromossomo.

O gene *WIFI*, candidato a associação com a característica de umbigo foi detectado em estudos de expressão, sendo relacionado com a quantidade de pelos produzidos em ratos. Esse gene que está presente no BTA 5 interage com proteínas morfogenética do osso envolvidas na formação e tamanho de folículos pilosos (Sharov et al., 2006).

Os genes *TMEM176A* e *TMEM176B* são expressos em células envolvidas com o sistema imune. Deficiência na transcrição desses genes está envolvida com maior incidência de inflamações da pele, como a psoríase (Drujont et al., 2016).

O *PKIB* foi detectado em estudos de associação em bovinos, o qual relaciona a expressão desse gene à aspectos adaptativos para regiões tropicais, uma vez que o

mesmo não foi identificado em indivíduos estudados fora destas regiões. Essa diferença foi apontada como resultante de pressões seletivas para sobrevivência de mamíferos (Amorim et al., 2015; Chung et al., 2009).

Na região apontada pelo GWAS que apresentou associação com pelagem, não foram relatados até o momento locos quantitativos relacionados a essa característica. No entanto, essa região pode ser candidata a QTL, por apresentar o gene *NTNG1* que foi associado à pelagem neste e em outros trabalhos, sugerindo a necessidade de mais estudos acerca dessa região.

O gene candidato *NTNG1*, foi confirmado como diferencialmente importante para a característica de pelagem lisa ou encaracolada em ovinos (Kang et al., 2013). A morfologia da pelagem possui importância para a adaptação do animal por afetar trocas térmicas de calor, atuando no isolamento térmico, na eficiência da termólise evaporativa (termorregulação que assegura perda de calor nos animais) e atributos termorreguladores associados (Barnabe, 1975; Li et al., 2014; Nicolau et al., 2004).

Nas regiões candidatas a associação com CE foram detectados QTLs e genes que desempenham função na taxa de prenhes das filhas, facilidade no parto e concepção por inseminação. Esses resultados corroboram com estudos anteriores que estabelecem a relação entre CE e caracteres de reprodução em fêmeas. A CE foi também relacionada com a capacidade reprodutiva de touros pela quantidade de esperma (Silva et al., 2001).

Há controvérsias entre a seleção de novilhas mediante resultados obtidos por meio de reprodutores, no entanto, estudos de correlação entre circunferência escrotal e características reprodutivas em fêmeas descendentes, resultaram em estimativas satisfatórias. Esses resultados demonstram a possibilidade de considerar a CE como indicador de caracteres reprodutivos desejáveis em fêmeas (filhas), como a idade à puberdade, que apresentou correlação de -0.70. Essa correlação negativa é desejável por

expressar que quanto maior a CE dos touros em menor idade suas filhas entrarão à puberdade (Santana et al., 2014).

Dentre os genes presentes em regiões apontadas pelo GWAS para possível associação com CE, mutações no gene *BRCA2* estão relacionadas com o desenvolvimento de câncer de mama, ovário e câncer de próstata em humanos (Fortuny et al., 2008).

*LOXLI* é um gene que codifica uma proteína envolvida na formação de fibras elásticas, sendo importante para estudos que buscam a compreensão de mecanismos relacionados com o prolapso de órgãos pélvicos. Em ratas e em primatas, mutações nesse gene acarretam em descida da vagina, do cólo do útero, da bexiga e prolapso retal através da descida da musculatura do assoalho pélvico (Consonni et al., 2012). Nos machos, mutações nesse gene foram associadas à diminuição da protrusão testicular, incluindo protrusão perineal, morfologia escrotal menos desenvolvida, redução da distância anogenital, testículos mal palpáveis ou não palpáveis e menor concentração espermática. Assim o *LOXLI* demonstrou, em estudos prévios, relação com a fertilidade em ratos (Wood et al., 2009).

O gene candidato *GZMB* está envolvido com o sucesso da gravidez em ratas. Foram relatados que alterações nesse gene acarretam em baixa fertilidade, isso ocorre devido a maior dificuldade de implantação do embrião no útero (Nuño-Ayala et al., 2012). A expressão do *GZMB* também foi identificada no endométrio de equinos e caracterizada como de grande importância no estabelecimento da gravidez (Tachibana et al., 2013).

Os genes candidatos a associação com o comprimento de umbigo, pelagem e CE foram submetidos à busca por relação com clusters. Para essa busca foram utilizados genes ortólogos mapeados em humanos. A utilização da informação de seres humanos

foi motivada por ser a espécie de mamíferos com genoma melhor anotado (Monaco et al., 2015).

Os genes *TMEM176A* e *TMEM176AB*, candidatos à associação com a característica comprimento de umbigo, são ortólogos em humanos e constituem o cluster *Androgen Induced 1* (AIG1, Id: GO:0005525). Genes pertencentes a esse cluster tem relação descrita com o câncer de próstata, por desempenharem importante função na regulação da apoptose e manutenção do ciclo celular (Levina et al., 2015). A provável causa de ausência de relação, entre os genes identificados para as características de pelagem e CE com clusters, é que os genes estejam atuando de maneira independente.

#### **4. Conclusão**

Nas janelas identificadas pelo GWAS foram detectados locos e genes candidatos previamente descritos na literatura, com função relacionada com as características estudadas, reforçando os indícios de associação com comprimento de umbigo, pelagem e circunferência escrotal. Esse conhecimento poderá contribuir para a compreensão da arquitetura genética envolvida na expressão desses caracteres. Além de colaborar com o desenvolvimento de modelos de predição de mérito genético para implementação de seleção genômica na raça Canchim, já que estas características fazem parte da avaliação genética tradicional, que é realizada anualmente pelo programa de melhoramento da raça.

#### **Declaração do Conselho de Ética**

Para esse estudo, os dados foram coletados de acordo com o comitê de ética de uso animal da Embrapa Pecuária Sudeste e da Embrapa Gado de Corte sob os números de protocolos 02/2009 e 10/2015, respectivamente.

## **Declaração de conflito de interesse**

Nenhum dos autores do manuscrito tem potenciais conflitos de interesse que devem ser divulgados.

## **Agradecimentos**

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão de bolsas de estudo; À Embrapa Gado de Corte e à Embrapa Pecuária Sudeste por disponibilizar os dados genotípicos utilizados. À Associação Brasileira de Criadores de Canchim (ABCCAN) e aos produtores por terem cedido gentilmente as amostras biológicas dos animais analisados. Ao Programa de Melhoramento Geneplus/Embrapa por disponibilizar as DEPs dos animais.

## **Referências Bibliográficas**

- Amorim, C.E.G., Daub, J.T., Salzano, F.M., Foll, M., Excoffier, L., 2015. Detection of convergent genome-wide signals of adaptation to tropical forests in humans. *PloS ONE* 10, e0121557.
- Barichello, F., Alencar, M.M.d., Torres Júnior, R.A.d.A., Silva, L.O.C.d., 2010. Herdabilidade e correlações quanto a peso, perímetro escrotal e escores visuais à desmama, em bovinos Canchim. *Pesqui. Agropec. Bras.*, 563-570.
- Barnabe, R.C., 1975. Variações estacionais no pelame de vacas da raça Jersey e sua correlação com a produção leiteira. *Braz. J. Vet. Res. Anim. Sci.* 12, 95-106.
- Bignardi, A., El Faro, L., Torres Júnior, R., Cardoso, V., Machado, P., Albuquerque, L.G.d., 2011. Random regression models using different functions to model test-day milk yield of Brazilian Holstein cows. *Genet. Mol. Res.*, 3565-3575.
- Buzanskas, M.E., Grossi, D.A., Ventura, R.V., Schenkel, F.S., Sargolzaei, M., Meirelles, S.L., Mokry, F.B., Higa, R.H., Mudadu, M.A., da Silva, M.V.B., 2014. Genome-wide association for growth traits in canchim beef cattle. *PloS ONE* 9, e94802.

- Cantor, R.M., Lange, K., Sinsheimer, J.S., 2010. Prioritizing GWAS results: a review of statistical methods and recommendations for their application. *A. J. Hum. Genet.* 86, 6-22.
- Cha, S.T., Talavera, D., Demir, E., Nath, A.K., Sierra-Honigmann, M.R., 2005. A method of isolation and culture of microvascular endothelial cells from mouse skin. *Microvasc. Res.* 70, 198-204.
- Chung, S., Furihata, M., Tamura, K., Uemura, M., Daigo, Y., Nasu, Y., Miki, T., Shuin, T., Fujioka, T., Nakamura, Y., 2009. Overexpressing PKIB in prostate cancer promotes its aggressiveness by linking between PKA and Akt pathways. *Oncogene* 28, 2849-2859.
- Clayton, D., 2012. snpStats: SnpMatrix and XSnpmatrix classes and methods. <http://www-gene.cimr.cam.ac.uk/clayton>
- Consonni, S.R., Werneck, C.C., Sobreira, D.R., Kühne, F., Moraes, S.G., Alvares, L.E., Joazeiro, P.P., 2012. Elastic fiber assembly in the adult mouse pubic symphysis during pregnancy and postpartum. *Biol. Reprod* 86, 151, 151-110.
- Corbet, N., Prayaga, K., Johnston, D., Burrow, H., 2007. Genetic variation in adaptive traits of cattle in North Australia. *Proc. Assoc. Advmt. Anim. Breed. Genet.* pp. 348-351.
- Dennis, G., Sherman, B.T., Hosack, D.A., Yang, J., Gao, W., Lane, H.C., Lempicki, R.A., 2003. DAVID: database for annotation, visualization, and integrated discovery. *Genome Biol.* 4, R60.
- Drujont, L., Lemoine, A., Moreau, A., Bienvenu, G., Lancien, M., Cens, T., Guillot, F., Bériou, G., Bouchet-Delbos, L., Fehling, H.J., 2016. ROR $\gamma$ t<sup>+</sup> cells selectively express redundant cation channels linked to the Golgi apparatus. *Sci. Rep* 6, 23682.
- Fernando, R., Garrick, D., 2008. GenSel-User manual for a portfolio of genomic selection related analyses. 3. Accessed: Jul. 14 2017. <http://taurus.ansci.iastate.edu/>

Fordyce, G., McGowan, M.R., Lisle, A., Muller, T., Allen, J., Duff, C., Holroyd, R.G., Corbet, N.J., Burns, B.M., 2014. Scrotal circumference of Australian beef bulls. *Theriogenology* 81, 805-812.

Fortuny, D., Balmaña, J., Graña, B., Torres, A., Cajal, T.R.y., Darder, E., Gadea, N., Velasco, A., López, C., Sanz, J., 2008. Opinion about reproductive decision making among individuals undergoing BRCA1/2 genetic testing in a multicentre Spanish cohort. *Hum. Reprod.* 24, 1000-1006.

Garrick, D.J., Taylor, J.F., Fernando, R.L., 2009. Deregressing estimated breeding values and weighting information for genomic regression analyses. *Genet. Sel. Evol.* 41, 1.

Gitton, Y., Dahmane, N., Baik, S., i Altaba, A.R., Neidhardt, L., Scholze, M., Herrmann, B.G., Kahlem, P., Benkahla, A., Schrunner, S., 2002. A gene expression map of human chromosome 21 orthologues in the mouse. *Nature* 420, 586-590.

Habier, D., Fernando, R.L., Kizilkaya, K., Garrick, D.J., 2011. Extension of the Bayesian alphabet for genomic selection. *BMC Bioinformatics* 12, 186.

Kang, X., Liu, G., Liu, Y., Xu, Q., Zhang, M., Fang, M., 2013. Transcriptome profile at different physiological stages reveals potential mode for curly fleece in Chinese tan sheep. *PloS ONE* 8, e71763.

Kinsella, R.J., Kähäri, A., Haider, S., Zamora, J., Proctor, G., Spudich, G., Almeida-King, J., Staines, D., Derwent, P., Kerhornou, A., 2011. Ensembl BioMarts: a hub for data retrieval across taxonomic space. *Database* 2011.

Koury Filho, W., Jubileu, J.d.S., Eler, J., Ferraz, J., Pereira, E., Cardoso, E., 2003. Parâmetros genéticos para escore de umbigo e características de produção em bovinos da raça Nelore. *Arq. Bras. Med. Vet. Zoot.* 55, 594-598.

Levina, E., Ji, H., Chen, M., Baig, M., Oliver, D., Ohouo, P., Lim, C.-u., Schools, G., Carmack, S., Ding, Y., 2015. Identification of novel genes that regulate androgen receptor signaling and growth of androgen-deprived prostate cancer cells. *Oncotarget* 6, 13088.

- Li, M., Tiirikka, T., Kantanen, J., 2014. A genome-wide scan study identifies a single nucleotide substitution in ASIP associated with white versus non-white coat-colour variation in sheep (*Ovis aries*). *Heredity* 112, 122-131.
- McGowan, M., Bertram, J., Fordyce, G., Fitzpatrick, L., Miller, R., Jayawardhana, G., Doogan, V., De Faveri, J., Holroyd, R., 2002. Bull selection and use in northern Australia: 1. Physical traits. *Anim. Reprod. Sci.* 71, 25-37.
- Mokry, F.B., Higa, R.H., de Alvarenga Mudadu, M., de Lima, A.O., Meirelles, S.L.C., da Silva, M.V.G.B., Cardoso, F.F., de Oliveira, M.M., Urbinati, I., Niciura, S.C.M., 2013. Genome-wide association study for backfat thickness in Canchim beef cattle using Random Forest approach. *BMC Genet.* 14, 47.
- Monaco, G., van Dam, S., Ribeiro, J.L.C.N., Larbi, A., de Magalhães, J.P., 2015. A comparison of human and mouse gene co-expression networks reveals conservation and divergence at the tissue, pathway and disease levels. *BMC Evol. Biol.* 15, 259.
- Nicolau, C., Silva, R., Mota, L., Veríssimo, C., 2004. Características da pele e do pelame em bovinos da raça Caracu. *Arch. Zootec.* 53, 25-34.
- Nuño-Ayala, M., Guillen, N., Arnal, C., de Martino, A., García-de-Jalón, J.-A., Gascón, S., Osaba, L., Osada, J., Navarro, M.-A., 2012. Cystathionine  $\beta$ -synthase deficiency causes infertility by impairing decidualization and gene expression networks in uterus implantation sites. *Physiol. Genomics*.
- Onteru, S.K., Gorbach, D.M., Young, J.M., Garrick, D.J., Dekkers, J.C., Rothschild, M.F., 2013. Whole genome association studies of residual feed intake and related traits in the pig. *PloS ONE* 8, e61756.
- Ortega, J., Daft, B., Assis, R., Kinde, H., Anthenill, L., Odani, J., Uzal, F.A., 2007. Infection of internal umbilical remnant in foals by *Clostridium sordellii*. *Vet. Pathol. Online* 44, 269-275.
- Porto-Neto, L.R., Reverter, A., Prayaga, K.C., Chan, E.K., Johnston, D.J., Hawken, R.J., Fordyce, G., Garcia, J.F., Sonstegard, T.S., Bolormaa, S., 2014. The genetic architecture of climatic adaptation of tropical cattle. *PloS One* 9, e113284.

R Development Core Team, 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. .

Saatchi, M., Garrick, D., Fernando, R., Boddicker, N., 2012. Comparison of Different Bayesian Methods for QTL Mapping in Hereford Beef Cattle Using 1 Mb Windows. Accessed: Jul. 17 2017. <https://pag.confex.com/pag/xx/webprogram/Paper3217.html>

Santana, M.H., Utsunomiya, Y.T., Neves, H.H., Gomes, R.C., Garcia, J.F., Fukumasu, H., Silva, S.L., Junior, G.A.O., Alexandre, P.A., Leme, P.R., 2014. Genome-wide association analysis of feed intake and residual feed intake in Nelore cattle. *BMC Genet.* 15, 1.

Santiago, G., Siqueira, F., Cardoso, F., Regitano, L., Ventura, R., Sollero, B., Souza, M., Mokry, F., Ferreira, A., Torres, R., 2017. Genomewide association study for production and meat quality traits in Canchim beef cattle. *J. Anim. Sci.*

Sargolzaei, M., Chesnais, J.P., Schenkel, F.S., 2014. A new approach for efficient genotype imputation using information from relatives. *BMC Genomics* 15, 478.

Sharov, A.A., Sharova, T.Y., Mardaryev, A.N., di Vignano, A.T., Atoyán, R., Weiner, L., Yang, S., Brissette, J.L., Dotto, G.P., Botchkarev, V.A., 2006. Bone morphogenetic protein signaling regulates the size of hair follicles and modulates the expression of cell cycle-associated genes. *Proc. Natl. Acad. Sci. U.S.A.* 103, 18166-18171.

Silva, R.G., La Scala Jr, N., Pocay, P.L.B., 2001. Transmissão de Radiação Ultravioleta Através do Pelame e da Epiderme de Bovinos1. *Rev. Bras. Zootec.* 30, 1939-1947.

Sollero, B., GULIAS-GOMES, C., Roso, V., Higa, R., YOKOO, M., Cardoso, L., Caetano, A., Cardoso, F., 2014. Goodness of fit comparisons among five bayesian models in genome-wide association of tick resistance in brazilian Hereford and Braford beef cattle. *Embrapa Pecuária Sul-Artigo em anais de congresso (ALICE)*. In: World congress on genetics applied to livestock production, 10., 2014, Vancouver, British Columbia, Canada. *Proceedings...* Champaign: ASAS, 2014.

Tachibana, Y., Nakano, Y., Nagaoka, K., Kikuchi, M., Nambo, Y., Haneda, S., Matsui, M., Miyake, Y.-i., Imakawa, K., 2013. Expression of endometrial immune-related genes possibly functioning during early pregnancy in the mare. *J. Reprod. Develop.* 59, 85-91.

Tsai-Morris, C.-H., Koh, E., Sheng, Y., Maeda, Y., Gutti, R., Namiki, M., Dufau, M.L., 2007. Polymorphism of the GRTH/DDX25 gene in normal and infertile Japanese men: a missense mutation associated with loss of GRTH phosphorylation. *Mol. Hum. Reprod.* 13, 887-892.

Tupy, O., Primavesi, O., Barbosa, P., 2006. Avaliação dos impactos econômicos, sociais e ambientais de tecnologias da Embrapa Pecuária Sudeste. 1. Utilização de touros da raça Canchim em cruzamento terminal com fêmeas da raça Nelore. Embrapa Pecuária Sudeste-Documentos (INFOTECA-E).

Utsunomiya, Y.T., Carmo, A.S., Neves, H.H., Carvalheiro, R., Matos, M.C., Zavarez, L.B., Ito, P.K., O'Brien, A.M.P., Sölkner, J., Porto-Neto, L.R., 2014. Genome-wide mapping of loci explaining variance in scrotal circumference in nellore cattle. *PLoS ONE* 9, e88561.

Windeyer, M., Leslie, K., Godden, S., Hodgins, D., Lissemore, K., LeBlanc, S., 2014. Factors associated with morbidity, mortality, and growth of dairy heifer calves up to 3 months of age. *Prev. Vet. Med.* 113, 231-240.

Wood, H.M., Lee, U.J., Vurbic, D., Sabanegh, E., Ross, J.H., Li, T., Damaser, M.S., 2009. Sexual Development and Fertility of *Lox11<sup>-/-</sup>* Male Mice. *J. Androl.* 30, 452-459.

Yu, M., Bell, R.H., Ho, M.M., Leung, G., Haegert, A., Carr, N., Shapiro, J., McElwee, K.J., 2012. Deficiency in nucleotide excision repair family gene activity, especially ERCC3, is associated with non-pigmented hair fiber growth. *PLoS ONE* 7, e34185.

## Anexo

**Anexo 1.** “Manhattan plot” dos efeitos padronizados do teste de associação genômica ampla por meio de Bayes B com  $\pi = 0,99$ , para as características de umbigo (A), pelagem (B) e circunferência escrotal (C) em bovinos da raça Canchim.

