

AURINO DE ARAUJO REGO NETO

**ESTRUTURA GENÉTICA E ASSOCIAÇÃO GENÔMICA AMPLA PARA
CARACTERÍSTICAS DE TAMANHO CORPORAL EM OVINOS DA RAÇA SANTA
INÊS**

TERESINA - PIAUÍ

2017

AURINO DE ARAUJO REGO NETO

**ESTRUTURA GENÉTICA E ASSOCIAÇÃO GENÔMICA AMPLA PARA
CARACTERÍSTICAS DE TAMANHO CORPORAL EM OVINOS DA RAÇA SANTA
INÊS**

Tese submetida ao Programa de Pós-Graduação em
Ciência Animal da Universidade Federal do Piauí como
requisito para obtenção do título de DOUTOR EM
CIÊNCIA ANIMAL.

Área de Concentração: Produção Animal.

Orientador: Prof. Dr. José Lindenberg Rocha Sarmiento

Co-orientador: Prof. Dr. Fábio Barros Britto

TERESINA - PIAUÍ

2017

FICHA CATALOGRÁFICA
Universidade Federal do Piauí
Biblioteca Setorial do Centro de Ciências Agrárias
Serviço de Processamento Técnico

R343e Rego Neto, Aurino de Araújo
Estrutura genética e associação genômica ampla para características de tamanho corporal em ovinos da raça Santa Inês / Aurino de Araújo Rego Neto - 2017.
89 f.: il.

Tese (Doutorado em Ciência Animal) – Universidade Federal do Piauí, Teresina, 2017.
Orientação: Prof. Lindenberg Rocha Sarmento

1. Gado ovinos deslanado 2. Homozigose 3. Parentesco real
I. Título

CDD 636.3

ESTRUTURA GENÉTICA E ASSOCIAÇÃO GENÔMICA AMPLA PARA
CARACTERÍSTICAS DE TAMANHO CORPORAL EM OVINOS DA RAÇA
SANTA INÊS

AURINO DE ARAUJO REGO NETO

Tese Aprovada em: 10/03/2017

Banca Examinadora:



Prof. Dr. José Lindenberg Rôcha Sarmiento (Presidente) / DZO/CCA/UFPI



Prof. Dr. José Elivalto Guimarães Campelo (Interno) / DZO/CCA/UFPI



Prof. Dr. Fábio Barros Britto (Interno) / CCN/UFPI



Prof. Dr. Luiz Antonio Silva-Figueiredo Filho (Externo) / IFMA



Prof. Dr. Natanael Pereira da Silva Santos (Interno) / CPCE/UFPI



Prof. Dr. Marcos Jacob de Oliveira Almeida (Externo) / EMBRAPA

Aos meus pais *Aldemar Ferreira dos Santos e Maria da Gloria Araujo Rego*, por todo amor, carinho, apoio e dedicação;
(*In memoriam*) aos meus avós, *Cinobelino José dos Santos e Joventina Ferreira dos Santos*, pelo afeto, respeito e amor.

Aos meus irmãos Irineu Araujo dos Santos, Aristeu Araujo dos Santo, Iricélia Araujo dos Santos e Petrônio Batista dos Santos

E demais familiares por contribuírem direta ou indiretamente para a realização dos meus sonhos.

(Dedico e Ofereço)

AGRADECIMENTOS

A Deus, pela plenitude da vida e fonte de toda sabedoria.

À Universidade Federal do Piauí, ao Programa de Pós-Graduação em Ciência Animal, pela oportunidade de conclusão de mais uma etapa da minha vida acadêmica,

A FAPEPI pela concessão da bolsa de estudo.

Ao professor Dr. José Lindenberg Rocha Sarmiento pela orientação, compreensão, incentivo, confiança, respeito e principalmente por toda amizade demonstrada. Minha eterna gratidão.

Ao professor Dr. Fábio Barros Britto pelas contribuições, auxílios, amizade e co-orientação.

Ao professor Dr. José Elivalto Guimarães Campelo pelas contribuições, auxílios e amizade.

Ao Grupo de Estudos de Genética e Melhoramento Animal (GEMA) pelo apoio e pelos momentos de prazer da convivência nas etapas boas e ruins desse trabalho.

A todos os meus familiares, avós, tios e primos que contribuíram muito com palavras de incentivo e foram importantes para a concretização deste trabalho.

A todos os amigos e colegas que de alguma forma estiveram presentes nesta caminhada, com uma palavra, gesto, apoio ou simplesmente por estarem presentes nos momentos de dificuldades enfrentados.

Muito obrigado!

LISTA DE TABELAS

CAPÍTULO I

- Tabela 1** Correlação entre coeficientes de endogamia genômicos e tradicional em ovinos da raça santa Inês ----- 49
- Tabela 2** Número de animais e tamanho dos segmentos de homozigose compartilhados por mais de 50% da população por cromossomos da raça Santa Inês.----- 50

CAPÍTULO II

- Tabela 1** Componentes de variância, estimativas de herdabilidade para características detamanho corporal em ovinos da raça Santa Inês----- 73
- Tabela 2** Identificação de genes com base na variância genética aditiva, explicada por janelas de 50 SNPs adjacentes para peso adulto de ovinos da raça Santa Inês-- 74
- Tabela 3** Identificação de genes com base na variância genética aditiva, explicada por janelas de 50 SNPs adjacentes para comprimento da perna de ovinos da raça Santa Inês----- 75
- Tabela 4** Identificação de genes com base na variância genética aditiva, explicada por janelas de 50 SNPs adjacentes para perímetro da perna de ovinos da raça Santa Inês----- 76
- Tabela 5** Identificação de genes com base na variância genética aditiva, explicada por janelas de 50 SNPs adjacentes para comprimento corporal de ovinos da raça Santa Inês----- 77
- Tabela 6** Identificação de genes com base na variância genética aditiva, explicada por janelas de 50 SNPs adjacentes para altura da cernelha de ovinos da raça Santa Inês----- 78

| | | |
|-----------------|---|----|
| Tabela 7 | Identificação de genes com base na variância genética aditiva, explicada por janelas de 50 SNPs adjacentes para de ovinos da raça Santa Inês----- | 79 |
|-----------------|---|----|

LISTA DE FIGURAS

CAPÍTULO I

| | | |
|-----------------|--|----|
| Figura 1 | Genotipagem com tecnologia <i>infinium da illumina</i> ----- | 51 |
| Figura 2 | Média de desequilíbrio de ligação (r^2) ao longo de diferentes distâncias (kb) entre marcadores em ovinos da raça Santa Inês----- | 51 |
| Figura 3 | Regressão entre parentesco genômico obtido pela matriz (G) com o parentesco tradicional baseado em <i>pedigree</i> obtido pela matriz (A) em animais da raça Santa Inês----- | 52 |
| Figura 4 | Histogramas do coeficiente de endogamia de ovinos da raça Santa Inês utilizando dados genômicos a partir dos segmentos de homozigose (FROH), homozigose esperada e observada (FHOM), matriz de parentesco genômica (FGRM) e registros de <i>pedigree</i> (FPED)----- | 53 |
| Figura 5 | Relação entre o número e tamanho médio dos segmentos de homozigose ROH em ovinos da Raça Santa Inês----- | 54 |
| Figura 6 | Número de regiões de homozigose compartilhadas entre indivíduos por cromossomos da raça Santa Inês----- | 54 |

CAPÍTULO II

| | | |
|-----------------|--|----|
| Figura 1 | Variância genética aditiva (Var) explicada por janelas de 50 SNPs adjacentes distribuídos por cromossomos (Chr) para peso a idade adulta de ovinos da raça Santa Inês.----- | 80 |
| Figura 2 | Variância genética aditiva (Var) explicada por janelas de 50 SNPs adjacentes distribuídos por cromossomos (Chr) para comprimento da perna de ovinos da raça Santa Inês.----- | 80 |

| | | |
|-----------------|--|----|
| Figura 3 | Variância genética aditiva (Var) explicada por janelas de 50 SNPs adjacentes distribuídos por cromossomos (Chr) para perímetro da perna de ovinos Santa Inês----- | 81 |
| Figura 4 | Variância genética aditiva (Var) explicada por janelas de 50 SNPs adjacentes distribuídos por cromossomos (Chr) para comprimento corporal de ovinos da raça Santa Inês----- | 81 |
| Figura 5 | Variância genética aditiva (Var) explicada por janelas de 50 SNPs adjacentes distribuídos por cromossomos (Chr) para altura da cernelha de ovinos da raça Santa Inês----- | 82 |
| Figura 6 | Variância genética aditiva (Var) explicada por janelas de 50 SNPs adjacentes distribuídos por cromossomos (Chr) para circunferência torácica em ovinos da raça Santa Inês----- | 82 |

REGO NETO, Aurino de Araujo. **Estrutura genética e associação genômica ampla para características de tamanho corporal em ovinos da raça Santa Inês**. 2017. 89f. Tese (Doutorado em Ciência Animal) – Universidade Federal do Piauí, Teresina, 2017.

O conhecimento da estrutura genética de uma população permite identificar fatores que podem interferir na eficiência do progresso genético pela seleção. Objetivou-se com este estudo avaliar a estrutura genética por meio de análise tradicional e genômica, bem como associações entre regiões cromossômicas com características de tamanho corporal de animais da raça Santa Inês utilizando metodologia de passo único (ssGWAS). Para tanto, foram utilizadas informações de 428 animais criados nos estados do Piauí e Maranhão com registro na Associação Brasileira de Criadores de Ovinos, dos quais foram genotipados 271, utilizando-se SNPs chip de alta densidade Ovine 50K da Illumina. Após o controle de qualidade, foram utilizados 51.874 SNPs. No Capítulo 1, foram utilizados dois métodos para calcular o parentesco entre os animais, matriz de numeradores do coeficiente de parentesco de Wright (A), com base na informação de *pedigree*, e a matriz de relacionamento genético genômico (G), com base nos marcadores SNPs. Quatro critérios foram utilizados e comparados para estimar a endogamia, informações *pedigree* (FPED); corridas de homozigose (FROH), frequência de homozigose observada e esperada (FHOM) e a matriz de relacionamento genômico (FGRM). Com a estatística r^2 estimou-se a extensão do desequilíbrio de ligação entre pares de SNPs adjacentes presentes nos cromossomos autossômicos. A identificação dos segmentos de homozigose (ROH) nos cromossomos autossômicos foi realizada considerando pelo menos 50 SNPs homozigotos em um segmento mínimo de 1.000 Kb por animal. O desequilíbrio de ligação (LD) com base na estimativa de r^2 foi de 0,4443. Os coeficientes de endogamia médios estimados FPED, FROH, FHOM, FGRM, foram 0,0004 0,035, 0,025, e 0,552, respectivamente. As estimativas de parentesco médio com base em A e G foram, respectivamente, 0,02 e 0,255. Foram identificados 4.022 seguimentos de homozigose no genoma, dos quais se destacam três regiões no cromossomo 16 que foram compartilhadas por mais de 50% da população, o que pode indicar a ocorrência de seleção intensa para características cuja expressão está regulada por genes localizados neste cromossomo. No Capítulo 2, para verificar a associação de marcadores SNPs com características de tamanho corporal, realizou-se análise de associação genômica ampla (GWAS) por meio da metodologia GWAS de passo único (ssGWAS) para estimar os efeitos de marcadores e associá-los às características peso a idade adulta (PA), comprimento corporal (CC), altura da cernelha (AC), circunferência torácica (CT), comprimento da perna (CP) e perímetro da perna (PP). Observaram-se associações com PA, CP e PP no cromossomo 6, com CT nos cromossomos 4, 7 e 13 e com AC e CC no cromossomo 4. As regiões identificadas neste estudo apresentam vários genes com conhecimento biológico descrito que poderão auxiliar na melhor compreensão da expressão destas características, assim como, poderão auxiliar nas tomadas de decisões em programas de seleção da raça Santa Inês.

Palavras chave: desequilíbrio de ligação, homozigose, ovinos deslanados, parentesco real.

REGO NETO, Aurino de Araujo. 2017. **Genetic structure and genome-wide association for body size traits in Santa Inês sheep** 89p. Thesis (Doctorate in Animal Science) – Federal University of the State of Piauí, Teresina, 2017.

Abstract: Knowledge on the genetic structure of a population allows for the identification of factors that may affect the efficiency of genetic progress caused by selection. In this study we aimed to evaluate the genetic structure of Santa Inês sheep, by means of traditional and genomic analyzes, and detect associations of cromosomal regions with body size traits using the single-step (ssGWAS) methodology. Therefore, information of 428 animals raised in the states of Piauí and Maranhão, and registered with the Brazilian Association of Sheep Breeders, was used. From those animals, 271 were genotyped using the high density Illumina Ovine SNP50K BeadChip. After quality control, 51,874 SNPs were used for genomic analysis. In chapter 1, the traditional genetic relationship matrix (A, based on pedigree information) and the genomic relationship matrix (G, based on SNP markers) were used to calculate kinship between animals. Four criteria were used and compared in order to estimate the inbreeding coefficient: *pedigree* information (FPED); runs of homozygosity (FROH); observed and expected frequency of homozygosity (FHOM); and the genomic relationship matrix (FGRM). The extent of linkage disequilibrium was estimated between all adjacent pairs of SNPs present in autosomal chromosomes by means of r^2 . Identification of runs of homozygosity (ROH) in autosomal chromosomes was performed by considering at least 50 homozygous SNPs in a minimal segment of 1,000 Kb by animal. Linkage disequilibrium (LD): $r^2=0.4443$. The estimated average inbreeding coefficients FPED, FROH, FHOM, and FGRM were 0.0004, 0.035, 0.025, and 0.552, respectively. The estimates of average relationship based on A and G were 0.02 and 0.255, respectively. A total of 4,022 runs of homozygosity were identified throughout the genome, where three regions on chromosome 16 stand out, because they were shared by more than 50% of the population. This is probably an indicative that intensive selection is occurring for traits of which the expression is controlled by genes located in chromosome 16. In chapter 2, a genome-wide association (GWA) analysis was performed using the single-step GWAS (ssGWAS) method, in order to estimate markers effects and associate them with adult body weight (PA), body length (CC), height at whithers (AC), thoracic circumference (CT), leg length (CP), and leg perimeter (PP). Associations with PA, CP, and PP (on chromosome 6), with CT (on chromosomes 4, 7 and 13), and with AC and CC (on chromosome 4) were observed. The regions identified in this study showed several genes biologically described which would help in the better understanding on the expression of the traits in study, and would help in decision-making in Santa Inês selection programs.

Keywords: hair sheep, homozygosity, linkage disequilibrium, realized kinship.

SUMÁRIO

| | |
|--|----|
| 1 INTRODUÇÃO | 14 |
| 2 REVISÃO DE LITERATURA..... | 15 |
| 2.1 Estrutura populacional..... | 15 |
| 2.2 Erros de <i>pedigree</i> | 16 |
| 2.3 Marcadores moleculares | 18 |
| 2.4 Segmentos de homozigose (ROH) | 20 |
| 2.5 Desequilíbrio de ligação | 22 |
| 2.6 Associação Genômica Ampla (GWAS) | 23 |
| 2.7 Referências Bibliográficas..... | 24 |
| 3 CAPÍTULO I..... | 30 |
| ESTRUTURA GENÉTICA E SEGMENTOS DE HOMOZIGOSE EM OVINOS DA RAÇA SANTA INÊS | |
| Resumo..... | 31 |
| Abstract | 32 |
| Introdução..... | 33 |
| Material e Métodos | 34 |
| Resultados e Discussão | 40 |
| Conclusões | 46 |
| Referências | 46 |
| 4 CAPÍTULO II..... | 55 |
| ASSOCIAÇÃO GENÔMICA AMPLA PARA CARACTERÍSTICAS DE TAMANHO CORPORAL EM OVINOS DA RAÇA SANTA INÊS | |
| Resumo..... | 56 |
| Abstract | 57 |
| Introdução..... | 58 |
| Material e Métodos | 59 |
| Resultados e Discussão | 63 |
| Conclusões | 69 |
| Referências | 69 |
| 5 CONSIDERAÇÕES GERAIS..... | 83 |
| 6 ANEXOS..... | 84 |

1 INTRODUÇÃO

Com a crescente popularidade da produção de carne na indústria ovina, geneticistas e criadores tem direcionado suas atenções para o melhoramento de características de crescimento e de produção de carne. Nesse sentido, o conhecimento das potencialidades e limitações das raças é importante, pois fornece orientação para programas de melhoramento genético e seleção.

No Brasil, a Santa Inês é uma das raças nativas que apresenta destaque. Segundo CARNEIRO et al. (2006), foi a raça com maior aceitação no país, pelos criadores de ovinos de corte. É considerada uma opção para a produção de carne em quase todas as regiões do Brasil, até para o semiárido, pois, apresenta potencial de adaptação às condições peculiares desse ambiente, além de apresentar excelente qualidade de pele e bom desenvolvimento ponderal, o que a coloca em posição estratégica como uma reserva de diversidade genética factível de uso em programas de melhoramento genético por meio da seleção.

Para que programas de melhoramento genético tornem-se viáveis, é imprescindível o conhecimento de fatores que podem interferir na eficiência do progresso genético por seleção, como por exemplo, o intervalo de geração, tamanho efetivo e a variabilidade genética dos rebanhos (MALHADO et al., 2008). Programas de seleção em populações com tamanho efetivo pequeno são limitados, tanto pelo pequeno número de animais disponível para reprodução, quanto pela necessidade de controlar a endogamia (FONTANESI et al., 2015).

As correlações de elevada magnitude entre características de tamanho corporal com características de crescimento e produção de carne indicam proporcionalidade durante a fase de crescimento entre o aumento do corpo do animal e o aumento da musculabilidade (SANTOS et al., 2016).

O controle genético de características de crescimento e produção de carne tem sido estudado sem identificar os genes ou variantes genéticas subjacentes à variação observada, com seleção fundamentada nas estimativas de valores genéticos calculados a partir de informações fenotípicas e de *pedigree* (GODDARD E HAYES, 2009).

Nesse sentido, os painéis de alta densidade de polimorfismo de base única (*Single Nucleotide Polymorphism – SNP*) surgiram como uma ferramenta essencial para os estudos de associação genômica ampla e seleção de características economicamente importantes (HERERRO-MEDRANO et al., 2013). De acordo com PETERS et al. (2012), estudos de associação entre SNPs provenientes de painéis de alta de densidade e características de interesse econômico possuem muitas aplicações práticas, por exemplo, detecção de QTLs e construção de bibliotecas de genes. Esses estudos são úteis para se conhecer a estrutura

genética da população e para posterior aplicação de seleção genômica e seleção assistida por marcadores.

Objetivou-se com esta pesquisa avaliar a estrutura genética por meio de análise tradicional e genômica, bem como a associação entre regiões cromossômicas com características de tamanho corporal em ovinos da raça ovinos Santa Inês, utilizando metodologia de passo único (ssGWAS). A pesquisa está estruturada conforme a Resolução 001/03 CCMCA, que estabelece as normas para elaboração e apresentação de Tese do Programa de Pós-Graduação em Ciência Animal da Universidade Federal do Piauí e os Capítulos I e II seguem normas de revista indexada da área (Revista Pesquisa Agropecuária Brasileira), onde serão submetidos à análise para posterior publicação.

O estudo foi segmentado a fim de melhorar a acomodação dos diferentes tópicos. Inicialmente, tem-se a revisão de literatura, que engloba os temas abordados nos Capítulos, aos quais servirá de referencial teórico, bem como as respectivas Referências Bibliográficas utilizadas. Segue com o Capítulo I, o qual trata sobre a Estrutura genética e segmentos de homozigose em ovinos da raça Santa Inês. O Capítulo II versa sobre Associação genômica ampla para características de tamanho corporal em ovinos Santa Inês. A finalização desta pesquisa é alcançada com as Considerações Finais.

2 REVISÃO DE LITERATURA

2.1. Estrutura genética

Uma população, do ponto de vista genético, compreende o conjunto de genes que se agrupam ao longo do tempo e que depende de muitos processos, dentre eles: seleção, mutação migração e deriva genética (FALCONER, 1987). O estudo da estrutura populacional é um importante critério para analisar a variabilidade genética inter e intrapopulacional balizada pela taxa de consanguinidade.

O conhecimento da estrutura populacional de uma raça é de suma importância em razão da necessidade de monitorar a diversidade genética, bem como para o controle de acasalamentos entre indivíduos aparentados, visto que, a perda da variabilidade genética pode comprometer o progresso genético futuro em características de importância econômica (FARIA et al., 2010). Barros et al. (2011) afirmaram que o estudo da estrutura genética de uma raça ou população permite conhecer como os genes estão sendo conduzidos ao longo das gerações, o que proporciona informações sobre o número de genes que lhe deram origem e estimam a participação destes na população atual.

As informações de *pedigree* foram utilizadas no estudo da estrutura genética da população da maioria das espécies de interesse econômico, como bovinos (GUTIÉRREZ et al., 2003), caprinos (BARROS et al., 2011) e ovinos (RODRIGUES et al., 2009; REGO NETO et al., 2014). Entretanto, a avaliação da população com base em informações genealógicas muitas vezes é prejudicada pela escassez e qualidade desse tipo de informação.

Com os avanços da genética molecular, a avaliação da estrutura genética de populações foi facilitada, pois os marcadores moleculares fornecem informações valiosas para o monitoramento da seleção artificial, podendo ser usados na identificação individual ou conjunto de animais que devem ser mantidos em processo de reprodução, apenas para evitar a perda de diversidade genética. Além disso, essas informações de parentesco entre os indivíduos podem ser utilizadas para evitar o acasalamento entre animais endogâmicos (DEASSIS et al., 2009).

Os marcadores SNPs permitem avaliar a heterozigosidade observada e esperada, bem como o conteúdo de informação polimórfica (PIC) e equilíbrio de Hardy-Weinberg EHW). Com a análise de variância molecular (AMOVA), além das estatísticas F (F_{IT} , F_{IS} e F_{ST}), que em conjunto descrevem a distribuição da diversidade genética na população e permitem inferir sobre a existência de estruturação populacional (GUTIÉRREZ et al., 2003).

Outro aspecto de interesse ao estudar uma população sob pressão de seleção é avaliar o nível de endogamia. O coeficiente de endogamia tradicional é estimado com base em dados de *pedigree*, os quais dependem da integridade e exatidão dos registros genealógicos disponíveis. No entanto, com as informações fornecidas por marcadores moleculares (painéis de chips SNP do genoma), pode-se estimar esse coeficiente com ou sem informações de *pedigree* (BOUQUET et al., 2011).

2.2 Erros de *pedigree*

Entende-se por seleção tudo aquilo que possa favorecer ou prejudicar determinados indivíduos, sob o ponto de vista da transmissão de seus genes às futuras gerações. Tradicionalmente, a seleção de animais realizada pelo fenótipo é considerada viável para a identificação de animais com alto potencial genético, o que corresponde a seleção genotípica, em que o animal é selecionado por meio da medida de seu potencial genético.

O sucesso do melhoramento genético animal depende essencialmente da adoção de métodos acurados de seleção, os quais demandam a predição dos valores genéticos dos

animais candidatos à reprodução. O desenvolvimento da tecnologia computacional e do método BLUP (*Best Linear Unbiased Prediction*, Melhor Predição Linear Não Viesada) associado ao modelo animal tornou-se o procedimento padrão de avaliação genética em bovinos de leite, bovinos de corte, ovinos e suínos (RESENDE e PEREZ, 1999).

O método BLUP foi desenvolvido por HENDERSON (1973). A utilização de informações dos animais e de seus parentes, considerando a correção dos fenótipos para os efeitos ambientais identificáveis, permite com o BLUP prever valores genéticos dos animais com acurácia (RESENDE e PEREZ, 1999).

O rigor da predição dos valores genéticos pode ser melhorado pela utilização de todas as informações de parentesco entre os animais que estão sendo avaliados. A metodologia de modelos mistos, para obtenção do BLUP, utiliza a matriz de numeradores do coeficiente de parentesco de Wright (A), que inclui todas as informações genealógicas disponíveis. O uso da matriz A faz com que a seleção com base nos valores genéticos preditos pelo BLUP apresente superioridade sobre outros métodos de seleção, especialmente quanto às características de baixa herdabilidade, em que o BLUP enfatiza as informações de parentes, em detrimento das informações do próprio indivíduo (HENDERSON, 1975). Todavia, quando as informações de *pedigree* contêm erros, a predição dos valores genéticos dos animais pode ser viesada e a seleção pouco eficiente.

A obtenção de informações corretas de *pedigree* é fundamental para um programa de melhoramento genético bem sucedido e sua importância aumentou com a introdução, nos anos 80, do modelo animal para a avaliação genética. No modelo animal, uma linhagem com erro, por exemplo, um pai de um animal identificado incorretamente, afetará a predição do valor genético do animal e de todos os seus parentes (VISSCHER et al., 2002).

Vale ressaltar que, mesmo quando o *pedigree* tem perfeito relacionamento entre indivíduos de uma população, o valor previsto de parentesco pode variar em consequência da segregação e recombinação dos genes (HILL e WEIR, 2011), o que não é considerado nos métodos tradicionais de predição de valor genético.

Os erros de *pedigree* reduzem o ganho genético esperado pela seleção com base nos valores genéticos preditos pelo BLUP, principalmente em características de baixa herdabilidade, nas quais a informação correta de parentesco é de grande importância (CARNEIRO et al., 2006). SENNEKE et al. (2004) concluíram que as estimativas dos efeitos genéticos diretos e maternos aumentam com a redução de erros de *pedigree* simulados. MORRISSEY et al. (2007), com a utilização de dados simulados, afirmaram que na ausência de efeitos maternos, o viés da herdabilidade cresceu com o aumento da taxa de erros no

pedigree e que, um erro de 10% no *pedigree* pode levar a reduções de 2 a 3% na resposta da seleção, em gado leiteiro, enquanto que, diferentes estudos relataram erros de *pedigree* observados em bovinos de até 23%, em vários países.

Adotando a matriz de parentesco genômica, VISSCHER et al. (2002) observaram taxa de erro de *pedigree* de 10%. Estes autores atribuíram que este valor tem impacto considerável no progresso genético do rebanho leiteiro no Reino Unido. BÉRÉÑOS et al. (2014) observam que erros de *pedigree* na matriz de parentesco provocaram viés de forma negativa nas estimativas de herdabilidade e covariâncias genéticas, e que é concebível que uma melhor informação de parentesco, seja por *pedigree* ou por métodos genômicos, pode melhorar as estimativas de herdabilidade e aumentar a acurácia das predições dos valores genéticos.

Em termos práticos, o aumento da acurácia leva ao aumento do ganho genético por promover o encurtamento do intervalo entre gerações, já que os animais jovens passam a ser utilizados como reprodutores precocemente, e também pelo aumento da acurácia em si, que apresenta relação proporcional com o ganho genético. REZENDE (2012) observou que a incorporação de informações genômicas nas avaliações genéticas de bovinos de corte aumentou a acurácia das estimativas de valores genéticos.

2.3 Marcadores moleculares

Os marcadores moleculares surgiram como ferramenta capaz de detectar variações genéticas no DNA e têm sido utilizados em análises genéticas em diferentes situações como, por exemplo, na caracterização de raças, identificação de clones, fluxo gênico, taxa de cruzamento, parentesco e construção de mapas genéticos. Os marcadores podem diferir em características importantes como abundância genômica, nível de polimorfismo detectável e informação genética, especificidade dos *loci* e reprodutibilidade (FERREIRA e GRATAPAGLIA, 1998).

O uso de informações obtidas a partir de técnicas de genética molecular no melhoramento genético animal tem como principal vantagem a utilização direta da informação do material genético na seleção, permitindo melhorias na eficiência de seleção e diminuição no tempo de avaliação dos animais, em comparação com a seleção tradicional com base essencialmente em dados fenotípicos. O uso de informações genéticas moleculares e fenotípicas para as avaliações caracteriza a seleção assistida por marcadores genéticos (DEKKERS, 2004).

Com os avanços tecnológicos, o desenvolvimento do sequenciamento automático possibilitou a detecção dos SNPs, que é um tipo de marcador molecular de alto desempenho, eficiência e baixo custo. Juntamente com as pequenas mutações de transições, onde ocorrem trocas de uma purina por outra purina (A/G), ou de uma pirimidina por outra pirimidina (C/T), e as transversões, que ocorrem quando há troca de uma purina por uma pirimidina, ou vice-versa (A/T ou C/G), os SNPs representam a forma mais frequente de variação genética natural, geralmente na maior parte dos organismos e, desta forma, estão sendo muito utilizados como uma nova geração de marcadores moleculares com diversas aplicações (CAIXETA et al., 2013).

A sigla SNP origina-se do inglês *Single Nucleotide Polymorphism* (polimorfismo de base única) e é utilizada para representar a forma mais comum de variação genética entre os indivíduos de determinada população. Os polimorfismos de base única estão distribuídos de forma abundante por todo o genoma, com alta variabilidade genética. Apresentam a vantagem de serem mais estáveis que os outros marcadores, porém pouco informativos por geralmente serem bialélicos (dois alelos por *locus*), o que requer a utilização de grande quantidade em análises genéticas. Outra vantagem é a possibilidade de automação, que permite a utilização de grande quantidade de indivíduos e de marcadores, assim como a genotipagem de alta velocidade (*ultra-high throughput genotyping*), na localização de *loci* de características quantitativas (QTL, *Quantitative Trait Loci*) (CAIXETA et al., 2013).

Os SNPs podem ocorrer em regiões codificadoras ou com função regulatória, bem como em espaços intergênicos, sem função determinada. Em regiões codificadoras, quando resultam em uma substituição de aminoácido na sequência proteica, são denominados não sinônimos, podendo a substituição ser conservativa ou não conservativa, em função das características dos aminoácidos envolvidos na troca. Nesses casos, pode haver modificações estruturais e funcionais na proteína (MEUWISSEN et al., 2001).

Embora SNPs sinônimos não alterem a sequência proteica, eles podem modificar a estrutura e a estabilidade do RNA mensageiro e, conseqüentemente, afetar a quantidade de proteína produzida, que pode ser afetada quando ocorrem alterações nas regiões não traduzidas do RNA mensageiro (5' UTR e 3' UTR). Além disso, polimorfismos gênicos podem promover processamentos alternativos, geração ou supressão de códons de terminação, alteração nos códons de iniciação da tradução e alterações no padrão de expressão de genes quando a troca de bases ocorre em sequências promotoras (GUIMARÃES e COSTA, 2002).

Plataformas para detecção de milhares a centenas de milhares dessas variações no DNA (Chips de SNPs de alta densidade) foram criadas e estão sendo utilizadas (PEREIRA et

al., 2013). Os chips de genotipagem de alta densidade permitem a genotipagem de milhares de SNPs distribuídos ao longo de todo o genoma da espécie em estudo, tornando possível, teoricamente, a partição de 100% da variabilidade genética aditiva de uma característica, propiciando a estimação do valor de substituição de alelos em cada um dos *loci* envolvidos com o fenótipo. Desta forma, depois de estabelecida a relação entre um conjunto de polimorfismos e uma característica de interesse, é possível estimar o valor genético de um indivíduo com base nos genótipos destes marcadores, ou seja, o valor genômico, sem a necessidade de avaliação fenotípica (MEUWISSEN et al., 2001).

O Consórcio Internacional do Genoma Ovino (*International Sheep Genomics Consortium*, ISGC, 2006) validou um painel de 60.000 marcadores SNP utilizando 2.819 animais de 74 raças de ovinos, em que o Brasil foi o único país da América do Sul a participar do Consórcio, com DNA de 98 animais pertencentes a três raças localmente adaptadas (Santa Inês, Morada Nova e animais Crioulos) (KIJAS et al., 2012).

No âmbito do projeto HapMap ovino, o Illumina OvineSNP50 BeadChip é utilizado para apresentar uma avaliação preliminar do LD em 74 raças diferentes de ovinos (RAADSMA, 2010). Um subgrupo de SNPs informativos a partir deste chip tem sido utilizado em ovinos selvagens para calcular a extensão do LD e avaliar a utilidade deste chip, que foi desenvolvido para ovinos domésticos e para a realização de estudos de associação do genoma em populações de ovinos selvagens (MILLER et al., 2010).

2.4 Segmentos de homozigose (ROH)

A disponibilidade de painéis de SNP de alta densidade permite identificar sequências longas e contínuas de genótipos homozigotos, definidas por KU et al. (2011) como segmentos de homozigose (*Runs of homozygosity*, ROH).

Segmentos de homozigose surgem quando os pais de um animal possuem um ancestral comum relativamente recente, isto é, quando pais são aparentados, estes compartilham grande parte dos seus genomas, que são idênticos por ascendência. Caso ambos os pais transmitam o mesmo segmento para sua progênie, ela será homozigota para o mesmo, dando origem aos ROH (KIM et al., 2013).

O tamanho de um segmento de homozigose contínuo depende do grau de parentesco dos progenitores, em que longos ROH são originários de endogamia recente, enquanto que ROH devido ao acasalamento de parentes distantes são geralmente menores (KIRIN et al., 2010).

Como os eventos de recombinação interrompem segmentos cromossômicos longos, ROH longos (~ 10 Mb) surgem como resultado de endogamia recente (até cinco gerações passadas), enquanto ROH mais curtos (~ 1 Mb) podem indicar efeito ancestral mais distante (até 50 gerações atrás), tais como efeito fundador da raça (HOWRIGAN et al., 2011). FISHER (1954) observou que o tamanho esperado de um segmento de DNA, que é idêntico por descendência, segue distribuição exponencial com média igual a $\frac{1}{2}g$ Morgans, em que “g” é o número de gerações passadas desde o ancestral comum. Logo, longos ROH (>5 Mb) refletem endogamia a partir de um ancestral comum que está a menos que 10 gerações passadas.

BROMAN e WEBER (1999) foram os primeiros a reconhecer que os longos trechos de homozigotos em humanos, muito provavelmente refletem autozigose, o que pode ter implicações de longo alcance para a saúde humana. Mapeamentos de homozigosidade utilizando dados genômicos foram realizados para identificar ROH em fenótipos complexos e associá-los com doenças, tais como esquizofrenia, transtorno bipolar, Parkinson, Alzheimer, câncer colorretal, de mama e próstata. Outras características como altura e peso também foram avaliadas usando análises de segmentos de homozigose (HOWRIGAN et al., 2011; YANG et al., 2012).

A base mais forte para o uso de ROH em genética de populações foi definida por MCQUILLAN et al. (2008), quando realizaram uma análise descritiva das populações europeias, incluindo ilhas Isoladas na Croácia e na Escócia, e definiram um novo coeficiente de endogamia genômica, calculado com base em segmentos de homozigose (FROH), que fornecem boa medida de autozigose individual de todo o genoma e permitem distinguir entre consanguinidade recente e antiga. Desde então, os ROH têm sido utilizados para estimar a endogamia (ZHANG et al., 2013). A extensão e frequência destes pode informar sobre a ascendência de um indivíduo em sua população. Hoje em dia, calculado a partir de ROH, o F de FROH é considerado o método que é mais poderoso para detectar efeitos de endogamia dentre as várias alternativas disponíveis, pois se correlaciona fortemente com carga mutação homozigótica (KELLER et al., 2011).

Em animais domésticos, SÖLKNER et al. (2010) e FERENČAKOVIĆ et al. (2011) foram os primeiros a estender o conceito de ROH, em trabalhos com bovinos, seguidos por PURFIELD et al. (2012) e FERENČAKOVIĆ et al. (2013). No entanto, o foco desses estudos tem sido a comparação entre coeficiente de endogamia calculado com base em segmentos de homozigose (FROH) e coeficiente de endogamia calculado com base em dados de *pedigree*

(FPED) em termos de comprimento de ROH, profundidade e qualidade de *pedigree*, algoritmos de cálculo e densidade do marcador.

Correlações de baixa magnitude entre as estimativas do coeficiente de endogamia obtidas a partir do *pedigree* e com base em dados de genotipagem de diferentes espécies foram observadas em alguns trabalhos (LOPES et al., 2013; SCRAGGS et al., 2013), o que pode ser explicado pelo efeito da variação da recombinação gênica, que não é considerada quando a endogamia é calculada a partir de dados de *pedigree*. Além disso, estimativas com base nos registros do *pedigree* negligenciam os efeitos da endogamia acumulada ao longo das gerações e não consideram os erros de *pedigree* acumulados entre as gerações passadas (MASTRANGELO et al., 2016).

2.5 Desequilíbrio de ligação (LD)

O desequilíbrio de ligação (LD) representa uma medida de correlação entre alelos em *loci* diferentes, representando a falta de independência, de forma estática, entre os alelos.

O LD existe entre dois *loci* ligados quando alelos presentes nestes ocorrem no mesmo haplótipo de maneira mais frequente do que o esperado na população. Estas associações alélicas são devidas principalmente à proximidade física, mas também são influenciados pela história da população e pelas forças evolutivas como, por exemplo, a recombinação entre os polimorfismos (KHATKAR et al., 2008). Além da taxa de recombinação, a magnitude e o decaimento do LD podem ser alterados por diversos eventos como miscigenação, deriva genética, efeito fundador, gargalos populacionais, mutação e seleção. Outros fatores, tais como idade da população, tamanho e taxa de crescimento populacional, migração e endocruzamento podem modificar os padrões de LD (O'BRIEN et al., 2014).

Em resposta à redução de tamanho da população provocada por seleção, deriva e /ou outros processos evolutivos, combinações específicas de alelos podem ser conservadas ao longo de muitas gerações mais frequentemente do que o esperado por acaso, conduzindo ao desenvolvimento de blocos de desequilíbrio de ligação (LD). O comprimento destes blocos genéticos co-herdados, conhecidos como haplótipos, é proporcional ao nível de LD em todo o genoma e seu estudo torna-se crucial para entender a dinâmica e características das populações, incluindo assinaturas de seleção, recombinação e gargalos (SÁNCHEZ-MOLANO et al., 2016).

Por interferir diretamente na dinâmica genética das populações, o LD deve ser entendido como um fenômeno genético. Seu efeito é visualizado na segregação não

independente dos alelos dos diferentes *loci* (HARTL e CLARK, 2010), resultando numa correlação entre eles durante a formação dos haplótipos (FLINT-GARCIA et al., 2003). Assim, sua existência acaba interferindo na formação dos genótipos dos indivíduos nas populações.

O êxito dos métodos de associação e seleção genômica depende da extensão do desequilíbrio de ligação em todo o genoma, o que pode variar entre populações. LD entre o marcador e um QTL (*Quantitative Trait Loci*) é apontado como um dos fatores que afetam a acurácia do valor genético genômico. Como um exemplo, a medida de desequilíbrio de ligação serve para avaliar o número de marcadores necessários para associar variação genética com características economicamente importantes (O'BRIEN et al., 2014).

A população com extenso LD exigirá uma densidade de marcador inferior. Em contraste, se o LD persistir ao longo de distâncias curtas, muito mais marcadores serão necessários para obter o mesmo poder para detectar associação (MEADOWS et al., 2008). Do mesmo modo, as assinaturas de regiões genômicas sob seleção positiva podem ser identificadas por meio do estudo da estrutura de bloco de haplótipo em todo o genoma (QANBARI et al., 2010).

2.6 Estudos de Associação Genômica Ampla - GWAS

Os avanços tecnológicos aliados ao fator custo-benefício possibilitaram a utilização de painéis densos de polimorfismos de nucleotídeo único (SNP), possibilitando o conhecimento de informações do genoma antes desconhecidas do melhoramento genético, o que tornou possível a investigação sobre a estrutura genética das características de interesse econômico. Sendo assim, estudos de associação genômica ampla (GWAS – *Genome Wide Association Studies*) estão sendo utilizados na busca por variantes genéticas associadas com a expressão fenotípica de diferentes características em diversas populações (WELLER e RON, 2011; VISSCHER et al., 2012; LIU et al., 2013; PONSUKSILI et al., 2014).

Os estudos de associação genômica ampla (GWAS) buscam combinar informações de genótipos, fenótipos e de *pedigree* dos indivíduos para identificar regiões do genoma que influenciam características de interesse. No cenário da produção animal, análises do tipo GWAS podem ser definidas pelo uso de painéis densos de SNPs para avaliar uma determinada população ou raça, onde as variações genéticas exercem função sobre qualquer característica fenotípica de importância econômica e têm o objetivo de revelar essas associações entre SNP e fenótipo de modo a identificar genes e/ou *loci* de características quantitativas (QTL – *Quantitative Trait Loci*) relacionados à expressão de fenótipos. Esses

estudos são úteis para se conhecer a estrutura genética da população e para posterior aplicação de seleção genômica e seleção assistida por marcadores (UTSUNOMIYA et al., 2013).

Segundo PETERS et al. (2012), estudos do tipo GWAS eram realizados pela metodologia de quadrados mínimos, aplicando-se correção de Bonferroni para se fazer inferências sobre QTLs a partir de efeitos individuais de SNPs, assumindo ausência de estrutura da população, o que, segundo SILVA et al. (2017), é muito improvável de acontecer quando se está trabalhando com dados reais. Com isso, os principais problemas enfrentados com essa metodologia são as altas taxas de falso-positivos e a superestimação dos efeitos de QTLs e dos SNPs.

Uma abordagem alternativa com a finalidade de corrigir tais problemas de GWAS foi proposta por WANG et al. (2012), onde todos os genótipos, fenótipos e informações de *pedigree* são considerados em uma única etapa (ssGWAS), permitindo assim, o uso de diferentes modelos e todas as relações de parentesco simultaneamente. Com esta abordagem, todos os SNPs são considerados ao mesmo tempo, junto com todos os fenótipos dos animais genotipados e não genotipados.

Trabalhos relacionados à associação genômica com características de crescimento, tamanho corporal e cor têm sido realizados em ovinos (KIJAS et al., 2012, ZHANG, et al., 2013, MATIKA et al., 2016).

REFERÊNCIAS BIBLIOGRÁFICAS

- BARROS, E.A.; RIBEIRO, M.N.; ALMEIDA, M.J.O.; ARAÚJO, A.M.³ Estrutura populacional e variabilidade genética da raça caprina Marota. **Archivos de Zootecnia**, v.60, p. 543-552, 2011.
- BÉRÉÑOS, C.; ELLIS, P. A.; PILKINGTON, J. G.; PEMBERTON, J. M. Estimating quantitative genetic parameters in wild populations: a comparison of *pedigree* and genomic approaches. **Molecular Ecology**, v. 23, p. 3434–3451, 2014.
- BOUQUET, A., VENOT, E., LALOÉ, D. FARABOSCO, F.; FOGH. A.; PABIU, T.; MOORE, K.; ERIKSSON, J.A.; RENAND, G.; PHOCOS, F. Genetic structure of the European Charolais and Limousin cattle metapopulations using *pedigree* analyses. **Journal of Animal Science**. v. 89, p. 1719-1730, 2011.
- BROMAN, K.W., WEBER, J. L. Long homozygous chromosomal segments in reference families from the centred' Étuded upolymor- phisme humain. **The American Journal of Human Genetics**. v. 65, p. 1493–1500, 1999.
- CAIXETA, E. T.; FERRÃO, L. F. V.; MACIEL- ZAMBOLIM, E. Marcadores moleculares. In: BORÉM, A.; FRITSCHÉ-NETO, R. **Biotecnologia aplicada ao melhoramento de plantas**: Suprema, 2013. cap. 2, p. 336.

- CARNEIRO, P.L.S.; MALHADO, C.H.M.; AZEVEDO, D.M.M.R.; SOUZA, L.G.R.; SILVA, D.C.; ALVES, A.A.; CAMPELO, J.E.G.; COSTA JÚNIOR, G.S. Estudo de populações de Ovinos Santa Inês utilizando técnicas de análise multivariada. **Revista Científica de Produção Animal**, v.8, n.1, 2006.
- DEASSIS, J. B.; DELAAT, D. M.; PEIXOTO, M.G.C.D.; BEREMANN, J.A.G.; FONSECA, C. G., CARVALHO, M. R.S. Genetic diversity and population structure in Brazilian Mangalarga Marchador horses. **Genetics and Molecular Research**, v. 8, p. 1519-1524, 2009.
- DEKKERS JC. Commercial application of marker- and gene-assisted selection in livestock: strategies and lessons. **Journal of Animal Science**, v. 82, p. E313–328, 2004.
- FALCONER, D.S. **Introdução à Genética Quantitativa**. Trad. M.A. Silva e J.C.Silva. Viçosa, MG: UFV, Impr. Univ, 1987. p. 279.
- FARIA, L.C. ; QUEIROZ, S. A. ; VOZZI, P. A. ; LÔBO, R. B. ; MAGNABOSCO, C. V. ; OLIVEIRA, J. A. Variabilidade genética da raça Brahman no Brasil detectada por meio de análise de *pedigree*. **Pesquisa Agropecuária Brasileira**, v.45, n.10, p.1133-1140, 2010.
- FERENČAKOVIĆ, M.; HAMZIC,E.; GREDLER,B.; CURIK, I.; SOLKNER, J. Runs of homozygosity reveal genome-wide autozygosity in the Austrian fleck- vieh cattle. **Agriculturae Conspectus Scientificus**. v.76, p. 325–328, 2011.
- FERENČAKOVIĆ, M.; HAMZIĆ, E.; GREDLER,B.; SOLBERG,T.R.; KLEMETSDAL,G.; CURIK, I.; SÖLKNER, J. Estimate of autozygosity derived from runs of homozygosity: empirical evidence from selected cattle populations. **Journal of Animal Breeding and Genetics**. v.130, p. 286–293, 2013.
- FERENČAKOVIĆ, M.; SOLKNER,J.; CURIK, I., Estimating autozygosity from high-throughput information: effects of SNP density and genotyping errors. **Genetics Selection Evolution**. v. 45, 2013.
- FERREIRA, M. E.; GRATTAPAGLIA, D. **Introdução ao uso de marcadores moleculares em análise genética**.3. ed. Brasília: Embrapa-Cernagen, 1998. 220p.
- FISHER, R. A. A fuller theory of “junctions” in inbreeding. **Heredity**, v. 8, p. 187–197,1954.
- FLINT-GARCIA, S. A.; THORNSBERRY, J. M.; BUCKLER, I.V. Structure of Linkage Disequilibrium in Plants. **Annual Review of Plant Biology**, Califórnia, v.54, n. 1, p. 357-374, 2003.
- FONTANESI, L.; SCOTTI, E.; SAMORÈ, A.B.; BAGNATO, A.; RUSSO, V. Association of 20 candidate gene markers with milk production and composition traits in sires of Reggiana breed, a local dairy cattle population. **Livestock Science**. v. 176, p. 14-21, 2015.
- GODDARD, M. E.; HAYES, B. J. Mapeamento de genes para traços complexos em animais domésticos e sua utilização em programas de melhoramento genético. **Nature Reviews Genetics**. v. 10, p. 381-391, 2009.

GUIMARÃES, P. E. M.; COSTA, M. C. R. SNPs: Sutis diferenças de um código. **Biociência** n. 26, p. 24-27, 2002.

GUTIÉRREZ, J.P.; ALTARRIBA, J.; DÍAZ, C.; QUINTANILLA, R.; CAÑÓN, J.; PIEDRAFITA, J. *Pedigree* analysis of eight Spanish beef cattle breeds. **Genetics Selection Evolution**. v. 63, p. 35-43, 2003.

HARTL, D. L.; A. G., CLARK. **Princípios de genética de populações**. In: Editora Artmed, 4. ed., 2010.

HERRERO-MEDRANO, J.; MEGENS, H-J.; GROENEN, M. A. M. RAMES, G.; BOSSE, M.; PÉREZ-ENCISO, M.; CROOIJMANS, R.P.M.A. Conservation genomic analysis of domestic and wild pig populations from the Iberian Peninsula. **BMC Genetics**, v. 14, p. 106-119, 2013.

HENDERSON, C.R. Sire evaluation and genetic trends. In: Animal Breeding and Genetics Symposium in Honor of J. Lush. **American Society of Animal Science**, v.3, p.10-41, 1973.

HENDERSON, C.R. Rapid method for computing the inverse of a relationship matrix. **Journal of Dairy Science**, v.58, p.1727-1730, 1975.

HILL, W.G.; WEIR, B.S. Variation in actual relationship as a consequence of Mendelian sampling and linkage. **Genetics Research**, v. 93, p. 47-64, 2011.

HOWRIGAN, D. P.; SIMONSON, M. A.; KELLER, M. C. Detecting autozygosity through runs of homozygosity: a comparison of three autozygosity detection algorithms. **BMC Genomics**. v. 12, p. 460-475, 2011.

KELLER, M.C.; VISSCHER. P.M.; GODDARD, M.E. Quantification of inbreeding due to distant ancestors and its detection using dense single nucleotide polymorphism data. **Genetics** v.189, 237-249, 2011.

KHATKAR, M. S.; NICHOLAS, F. W.; COLLINS, A. R.; ZENGER, K.R.; CAVANAGH, J. A.L.; BARRIS, W.; SCHNABEL, R.D.; TAYLOR, J.F.; RAADSMA, H.W. Extent of genome-wide linkage disequilibrium in Australian Holstein-Friesian cattle based on a high-density SNP panel. **BMC Genomics**, v. 9, p.187-205, 2008.

KIJAS, J.W.; LENSTRA, J.A.; HAYES, B.; BOITARD, S.; PORTO NETO, L. R.; SAN CRISTOBAL, M.; SERVIN, B.; MCCULLOCH, R.; WHAN, V.; GLEITZEN, K. PAIVA, S. BARENDSE, W. et al. Genome-Wide Analysis of the World's Sheep Breeds Reveals High Levels of Historic Mixture and Strong Recent Selection. **PLoS Biology**, v.10, 2012.

KIM, E.S.; COLE, J.B.; HUSON, H.; WIGGANS, G.R.; VANTASSELL, C.P.; Crooker, B.A.; Liu, G.; Da, Y.; Sonstegard, T.S. Effect of artificial selection on runs of homozygosity in U.S. Holstein cattle. **PLoS One**, v. 8, 2013.

KIRIN, M.; MCQUILLAN, R.; FRANKLIN, C.; CAMPBELL, H.; MCKEIGUE, P.; WILSON, J. Genomic runs of homozygosity record population history and consanguinity. **PLoS One**, v. 5, 2010.

- KU, C.S.; NAIDOO, N.; TEO, S.M.; PAWITAN, Y. Regions of homozygosity and their impact on complex diseases and traits. **Human Genetics**. v. 129, p. 01-15, 2011.
- LIU, R.; SUN, Y.; ZHAO, G. et al. Genome-Wide Association Study Identifies Loci and Candidate Genes for Body Composition and Meat Quality Traits in Beijing-You Chickens. **Plos One**, v.8, 2013.
- LOPES, M. S.; SILVA, F. F.; HARLIZIUS, B.; DUIJVESTELIJN, N.; LOPES, P. S. GUIMARÃES, S. E. F.; KNOL, E. F. Improved estimation of inbreeding and kinship in pigs using optimized SNP panels. **BMC Genetics**, v. 14, 2013.
- MATIKA, O.; RIGGIO, V.; ANSELME-MOIZAN, M.; LAW, A.S.; POW-WONG, R.; ARCHIBALD, A.L.; BISHOP, S.C. Association genome-wide reveals QTL for growth, bone and in vivo carcass traits as assessed by computed tomography in Scottish Black sheep. **Genetics Selection Evolution**. v. 48, p. 11-26, 2016.
- MALHADO, C. H. M.; CARNEIRO, P.L.S.; PEREIRA, D.G.; MARTINS FILHO, R. Progresso genético e estrutura populacional do rebanho Nelore no Estado da Bahia. **Pesquisa Agropecuária Brasileira**, v.43, n.9, p.1163-1169, 2008.
- MASTRANGELO, S., TOLONE, M., GERLANDO R.; FONTANESI, L.; SARDINA, M.T.; PORTALANO, B. Genomic inbreeding estimation in small populations: evaluation of runs of homozygosity in three local dairy cattle breeds. **Animal**, v. 5, p. 1-9, 2016.
- MCQUILLAN, R.; LEUTENEGGER, A.; ABDEL-RAHMAN, R. et al. Runso fhomozygosity in European populations. **American Journal of Human Genetics**. v. 83, p. 359–372, 2008.
- MEADOWS, J.R.; CHAN, E.K.; KIJAS, J.W. Linkage disequilibrium compared between five populations of domestic sheep. **BMC Genet**, v. 9, p. 61-71, 2008.
- MEUWISSEN, T. H. E.; HAYES, B. J.; GODDARD, M. E. Prediction of total genetic value using genome wide dense marker maps. **Genetics**, v. 157, p. 1819-1829, 2001.
- MILLER, J.M.; POISSANT, J.; KIJAS, J.W.; COLTMAN, D.W. International Sheep Genomics Consortium: A genome-wide set of SNPs detects population substructure and long range linkage disequilibrium in wild sheep. **Molecular Ecology Resources**, v. 11, p. 314–322. 2010.
- MORRISSEY, M.B.; WILSON, A.J.; PEMBERTON, J.M.; FERGUSON, M.M. A framework for power and sensitivity analyses for quantitative genetic studies of natural populations, and case studies in Soay sheep (*Ovis aries*). **Journal of Evolutionary Biology**, v. 20, p. 2309–2321, 2007.
- O'BRIEN, A. M. P.; MÉSZÁROS, G.; UTSUNOMIYA, Y. T.; SONSTERGARD, T. S.; GARCIA, F. J.; TESSELL, C. P. V.; CAVALHEIRO, R.; SILVA, M. V. B.; SOLKNER, J. Linkage disequilibrium levels in *Bos indicus* and *Bos taurus* cattle using medium and high density SNP chip data and different minor allele frequency distributions. **Livestock Science**, v. 166, p. 121-132, 2014.

PEREIRA G.L., ROSA K.O., CURI R.A., REGITANO L.C.A. & MOTA M.D.S. Estado da arte do seqüenciamento genômico na pecuária. **Ars Veterinaria**, v. 29, n. 3, p. 190-199, 2013.

PETERS, S. O.; KIZILKAYA, K.; GARRICK, D. J.; FERNANDO, R. L.; REECY, J. M.; WEABER, R. L.; SILVER, G. A.; THOMAS, M. G. Bayesian genome wide association analyses of growth and yearling ultrasound measures of carcass traits in Brangus heifers. **Journal of Animal Science**, v. 90, p. 3398-3409, 2012.

PONSUKSILI, S.; MURANI, E.; TRAKOOLJUL, N.; SCHWERIN, M.; WIMMERS, K. Discovery of Candidate Genes for Muscle Traits Based on GWAS Supported by e QTL analysis. **International Journal of Biological Sciences**, v.10, p.327-337, 2014.

PURFIELD, D.C.; BERRY, D.P.; MCPARLAND, S.; BRADLEY, D.G. Runs of homozygosity and population history in cattle. **BMC Genetics**, v. 13, p. 70-81, 2012.

QANBARI, S.; PIMENTEL, E.C.; TETENS, J.; THALLER, G.; LICHTNER, P.; SHARIFI, A.R.; SIMIANER, H. The pattern of linkage disequilibrium in German Holstein cattle. **Animal Genetics**, v. 41, p. 346–356, 2010.

RAADSMA, H.W: International Sheep Genomics Consortium (ISGC): Linkage disequilibrium in the sheep genome: Findings from the ISGC HapMap initiative. In **Proceedings of the XVIII Plant and Animal Genome (PAG)**. San Diego, CA, USA: 2010.

REGO NETO, A. de A.; SARMENTO, J.L.R.; SANTOS, N. P. S.; BIAGIOTTI, D.; SANTOS, G.V.et al. Estrutura e distribuição geográfica do rebanho de ovinos Santa Inês no Estado do Piauí. **Revista Brasileira Saúde e Produção Animal**, v. 15, n. 2, p. 272-280, 2014.

RESENDE, M. D. V., PEREZ, J.R. H.R. Melhoramento animal: predição de valores genéticos pelo modelo animal – BLUP em bovinos de leite, bovinos de corte, ovinos e suínos. **Archives of Veterinary Science**, v. 4, n. 1, p.17-29, 1999.

REZENDE, M. F. **Incorporação de marcadores genéticos em programas de melhoramento genético de bovinos de corte**. 2012. 88 f. Tese (Faculdade de zootecnia e engenharia de alimentos), Universidade de São Paulo, São Paulo, 2012.

RODRIGUES, D. S. ; RIBEIRO, M.N. ; OLIVEIRA, S.M.P. ; MELO LIMA, F.A. ; VILLARROEL, A.B.S. et al. Estrutura populacional de um rebanho da Raça Morada nova como contribuição para a conservação. **Ciência Animal**, v.19, n.1, p.103-110, 2009.

SÁNCHEZ-MOLANO, E.; TSIOKOS, D.; CHATZIPLIS D. A practical approach to detect ancestral haplotypes in livestock populations. **BMC Genetics**, v.17, p. 91-102, 2016.

SCRAGGS, E.; ZANELLA, R.; WOJTOWICZ, A.; TAYLOR, J.F.; GASKINS, C.T.; REEVES, J.J.; de AVILA, J.M.; NEIBERGS, H.L. Estimation of inbreeding and effective population size of full-blood wagyu cattle registered with the American Wagyu Cattle Association. **Journal of Animal Breeding and Genetics**, v. 131, n. 1, p.3-10, 2013.

SENNEKE, S.L.; MACNEIL, M.D.; VAN VLECK, L.D. Effects of sire misidentification on estimates of genetic parameters for birth and weaning weights in Hereford cattle. **Journal of Animal Science**, v. 82, p. 2307–2312, 2004.

SILVA R, M.O.; STAFUZZA, N.B.; FRAGOMENI, B.O.; CAMARGO, G.M.F.; CEACERO, T.M.; et al. Genome-Wide Association Study for Carcass Traits in an Experimental Nelore Cattle Population. **PLoS ONE**. v. 12, 2017.

SÖLKNER, J.; FRKONJA, A.; RAADSMA, H.W.; JONAS, E.; THALLER, G.; GOOTWINE, E.; SEROUSSI, E.; FUERST, C.; EGGER-DANNER, C.; GREDLER. Estimation of Individual Levels of Admixture in Crossbred Populations from SNP Chip Data: Examples with Sheep and Cattle Populations. **Interbull Bulletin**, n. 42, p. 62- 66, 2010.

UTSUNOMIYA, Y.T.; CARMO, A.S.; CARVALHEIRO, R. et al. Genome-wide association study for birth weight in Nelore cattle points to previously described orthologous genes affecting human and bovine height. **BMC Genetics**, v.14, 2013.

VISSCHER, P. M.; BROWN, M. A.; MCCARTHY, M. I.; YANG, J. Five Years of GWAS Discovery. **The American Journal of Human Genetics**, v.90, p.7-24, 2012.

WANG, S. J.; LIU, W. J.; SARGENT, C. A.; ZHAO, S. H.; LIU, H. B.; LIU, X. D.; WANG, C.; HUA, G. H.; YANG, L. G.; AFFARA, N. A.; ZHANG, S. J. Effects of the polymorphisms of Mx1, BAT2 and CXCL12 genes on immunological traits in pigs. **Molecular Biology Reports**, v. 39, p. 2417-2427, 2012.

WELLER, J. I.; RON, M. Invited review: Quantitative trait nucleotide determination in the era of genomic selection. **Journal of Dairy Science**, v.94, 2011.

YANG, H. C.; CHANG, L. C.; LIANG, Y. J.; LIN, C. H.; WANG, P. L. A Genome-wide Homozygosity Association Study Identifies Runs of Homozygosity Associated with Rheumatoid Arthritis in the Human Major Histocompatibility Complex. **Plos One**, v. 7, 2012.

ZHANG, L.; LIU, J.; ZHAO, F.; REN, H.; XU, L.; LU, J.; et al. Genome-Wide Association Studies for Growth and Meat Production Traits in Sheep. **Plos One**, v. 8, 2013.

3 CAPÍTULO I

ESTRUTURA GENÉTICA E SEGUIMENTOS DE HOMOZIGOSE EM OVINOS DA RAÇA SANTA INÊS

Estrutura genética e segmentos de homozigose em ovinos da raça Santa Inês

Aurino de Araujo Rego Neto⁽¹⁾, José Lindenberg Rocha Sarmento⁽²⁾

⁽¹⁾Pós Graduando do Programa de Pós-Graduação em Ciência Animal – UFPI. e-mail: aurinorego@hotmail.com

⁽²⁾Professor da UFPI/DZO e Bolsista do CNPq, e-mail: sarmento@ufpi.edu.br

Resumo: Objetivou-se com este estudo avaliar a estrutura genética da raça Santa Inês por meio de análise tradicional e genômica. Foram utilizadas informações de 428 animais criados nos estados do Piauí e Maranhão, dos quais foram genotipados 271, utilizando-se SNPs Chip de alta densidade Ovine 50K da Illumina. Após o controle de qualidade, foram utilizados 51.874 SNPs. Foram utilizados dois métodos para calcular o parentesco entre os animais, matriz de relacionamento genético tradicional (A), com base na informação de *pedigree*, e a matriz de relacionamento genético genômico (G), fundamentada nos marcadores SNP's. Quatro critérios foram utilizados e comparados para estimar a endogamia, informações *pedigree* (FPED); seguimentos de homozigose (FROH), frequência de homozigose observada e esperada (FHOM) e a matriz de relacionamento genômico (FGRM). Com a estatística r^2 estimou-se a extensão do desequilíbrio de ligação entre pares de SNPs adjacentes presentes nos cromossomos autossômicos. A identificação dos segmentos de homozigose (ROH) nos cromossomos autossômicos foi realizada considerando pelo menos 50 SNPs homozigotos em um segmento mínimo de 1.000 Kb por animal. O desequilíbrio de ligação (LD) r^2 foi de 0,4443. Os coeficientes de endogamia médios estimados FPED, FROH, FHOM, FGRM, foram 0,0004 0,035, 0,025, e 0,552 respectivamente. As estimativas de parentesco médio com base em A e G foram, respectivamente, 0,02 e 0,255. Foram identificados 4.022 seguimentos de homozigose no genoma, dos quais se destacam três regiões no cromossomo 16 que foram compartilhadas por mais de 50% da população, o que pode indicar a ocorrência de seleção

intensa para características cujo a expressão está regulada por genes localizados neste cromossomo.

Palavras chave: endogamia, ovinos deslanados, *pedigree*, segmentos de homozigose.

Genetic structure and runs of homozygosity in Santa Inês sheep

Abstract: In this study we aimed to evaluate the genetic structure of the Santa Inês breed by using traditional and genomic analyzes. We used information of 428 animals raised in the states of Piauí and Maranhão, and registered with the Brazilian Association of Sheep Breeders. From those animals, 271 were genotyped using the high density Illumina Ovine SNP50K BeadChip. After quality control, 51,874 SNPs were used for genomic analysis. The traditional genetic relationship matrix (A), based on *pedigree* information, and the genomic relationship matrix (G), based on SNP markers, were used to calculate relationship between animals. In order to estimate the inbreeding coefficient, four criteria were used and compared: pedigree information (FPED); runs of homozygosity (FROH); observed and expected frequency of homozygosity (FHOM); and the genomic relationship matrix (FGRM). The extent of linkage disequilibrium was estimated between all adjacent pairs of SNPs present in autosomal chromosomes by means of r^2 . Identification of runs of homozygosity (ROH) in autosomal chromosomes was performed by considering at least 50 homozygous SNPs in a minimal segment of 1,000 Kb by animal. Linkage disequilibrium (LD): $r^2=0.4443$. The estimated average inbreeding coefficients FPED, FROH, FHOM, and FGRM were 0.0004, 0.035, 0.025, and 0.552, respectively. The estimates of average relationship based on A and G were 0.02 and 0.255, respectively. A total of 4,022 runs of homozygosity were identified throughout the genome, where three regions in chromosome 16 stand out, because they were

shared by more than 50% of the population. This is probably an indicative that intensive selection is occurring for traits of which the expression is controlled by genes located in chromosome 16.

Keywords: hair sheep, inbreeding, pedigree, runs of homozygous.

Introdução

O conhecimento da estrutura genética em uma população é importante, não só como guia para as intervenções genéticas, mas para avaliar os resultados da seleção. Informações da estrutura da população combinadas com informações sobre mudanças genéticas na população podem auxiliar programas de melhoramento genético e conservação. Tradicionalmente feitos com base em informações de *pedigree*, os cálculos dos parâmetros populacionais podem ser subestimados, pois dependem da integridade e exatidão dos registros genealógicos disponíveis, além de muitas vezes a genealogia ser complexa e ter laços potencialmente desconhecidos, como também o efeito da variação da recombinação gênica, que não é considerado.

Devido à disponibilidade de painéis de SNP de alta densidade associada com o avanço das técnicas utilizadas no melhoramento genético animal, tornou-se possível a utilização de informações mais precisas sobre os alelos idênticos por descendência (IBD), por possuírem um ancestral comum no genoma de um indivíduo, e alelos idênticos independentemente de serem ou não herdados de um antepassado recente (IBS), que podem ser compartilhados por ancestrais comuns, ausentes no *pedigree*, tornando possível a utilização da matriz de parentesco genômica denominada *Genomic Relationship Matrix* (GRM) (Forni et al., 2011). O que contribui de forma significativa para obtenção de estimativas de parâmetros populacionais mais acurados.

A genotipagem com painéis de alta densidade permitem avaliar as similaridades no padrão do desequilíbrio de ligação (LD, *Linkage Disequilibrium*) em toda a população, fornecendo informações sobre o parentesco da mesma (Wang et al., 2013). Assim como o LD, os segmentos de homozigose (ROH- *Runs of Homozygosity*) apresentam indicativo sobre a estrutura da população, fornecendo informações sobre o histórico demográfico e eventos de endogamia da mesma (Bosse et al., 2012; Purfield et al., 2012).

Mapeamentos de homozigose utilizando dados genômicos foram realizados para identificar os ROH em fenótipos complexos e associá-los com doenças, tais como esquizofrenia, transtorno bipolar, Parkinson, Alzheimer, câncer colorretal, de mama e próstata. Outras características como altura e peso foram avaliadas usando análises de segmentos de homozigose, tendo sido identificados segmentos de homozigose associados com a doença de Alzheimer (Howrigan et al., 2011; Yang et al., 2012).

Tendo em vista a importância do conhecimento da estrutura genética de uma população para programas de melhoramento e conservação, objetivou-se com esta pesquisa avaliar a estrutura genética da população de ovinos da raça Santa Inês por meio de: (i) análises de desequilíbrio de ligação; (ii) estimação dos coeficientes de parentesco e endogamia utilizando dados genômicos e de *pedigree*; (iii) análise do número e tamanho de segmentos de homozigose ROH.

Material e Métodos

Dados

A pesquisa foi realizada com informações de 428 ovinos da raça Santa Inês, pertencentes a seis criatórios localizados na região Meio-norte do Brasil, dos quais 271 foram genotipados.

Extração de DNA

A extração do DNA foi realizada com o Kit *Qiagen DNeasy Blood & Tissue*, segundo o protocolo estabelecido pelo fabricante. Terminada a extração, a qualidade do DNA foi verificada em gel de agarose a 1% em tampão SB 1X por 60 minutos a 60V. Os géis foram posteriormente visualizados e registrados em fotodocumentador (*L-PIX Loccus Biotecnologia*). O DNA obtido foi estocado em freezer a -20°C até as análises subsequentes. Após a extração, foi realizado a quantificação do DNA, para isso foi utilizado o Fluorespectofotômetro NanoDrop 3300 com o kit intercalante AccueBlue dsDNA, que estima a concentração de DNA das amostras, essencial para preparação para genotipagem.

Genotipagem

A genotipagem foi realizada utilizando o chip de alta densidade *BeadChip OvineSNP50k* contendo 54.241 SNPs uniformemente espaçados nos cromossomos. O chip de SNPs foi fabricado de acordo com a tecnologia Infinium da Illumina™, usando a plataforma iScan. O protocolo de genotipagem foi seguido de acordo com o que foi estabelecido pelo fabricante (disponível em www.illumina.com). Este consistiu nos seguintes passos principais: (1) amplificação isotérmica e (2) incubação das amostras (*overnight*); (3) fragmentação das amostras, (4) precipitação e resuspensão, (5) preparação do *BeadChip* e (6) hibridização (*overnight*); (7) extensão enzimática de uma única base, (8) visualização das imagens na plataforma *iScan* e (9) resultados gerados (Figura 1).

O processo de amplificação isotérmica visou aumentar à ordem de centenas de vezes a concentração do DNA, evitando assim, que esse fosse um fator limitante e influenciasse de forma negativa os resultados. O material sofreu precipitação com álcool e resuspensão. O DNA foi submetido à ação de enzimas do tipo *end-point* que fragmentam em locais específicos da sequência, evitando a quebra excessiva do DNA. Posteriormente a esses

processos, o DNA genômico se anelou a 50 bases ligadas covalentemente a cada *Bead*, o qual representa um *locus* específico. A lâmina foi submetida a corantes que foram incorporados aos *Beads* e que permitiram a diferenciação dos genótipos. Por fim, os fluoróforos ligados aos *Beads* foram submetidos a um conjunto de *laser* que permitiu que o sinal de fluorescência fosse captado e analisado pela plataforma *iScan*.

Controle de qualidade e análise genética

O controle de qualidade foi realizado utilizando o programa PLINK v.1.9 (Purcell et al., 2007) para remoção de marcadores que falharam em mais de 5% das amostras e amostras que falharam em mais de 5% dos marcadores. Além disso, os registros dos SNPs dos cromossomos sexuais foram excluídos, porque a recombinação nestes cromossomos tende a ocorrer de forma diferente dos autossômicos. Nenhum filtro adicional foi considerado no estudo, uma vez que a remoção de alelos raros em análises de estrutura de população pode levar a obtenção de resultados superestimados ou subestimados quando se trabalha com amostras de populações (Bosse et al., 2012).

Após o controle de qualidade, foram calculadas estimativas de Heterozigosidade Esperada (H_e), Heterozigosidade Observada (H_o), segmentos de homozigose (ROH), coeficientes de endogamia calculada pela frequência de homozigose observada e esperada (FHOM), e coeficiente de endogamia calculado pelas regiões de homozigose (FROH), coeficiente de endogamia calculado pela matriz de relacionamento genético (FGRM) e desequilíbrio de ligação (LD).

O desequilíbrio de ligação entre dois SNPs foi avaliado utilizando o r^2 descrito por (Hill & Robertson, 1968) pela fórmula:

$$r^2 = \frac{D^2}{f(A)f(a)f(B)f(b)}$$

em que $D = f(AB) - f(A)f(B)$ e $f(AB), f(A), f(B), f(b)$ são as frequências observadas do haplótipo AB e dos alelos A, a, B, e b, respectivamente. Se os dois *loci* são independentes, então a frequência esperada do haplótipo *Ab* $f(AB)$ é calculada como o produto entre $f(A)$ e $f(B)$. Se $f(AB)$ é maior ou menor do que o esperado, então esses dois *loci* em particular tendem a segregar conjuntamente e estão em LD. As medidas de r^2 foram calculadas utilizando o programa computacional PLINK v.1.9 (Purcell et al., 2007).

A identificação dos segmentos de homozigose ROH foi realizada com auxílio do programa computacional PLINK v.1.9 (Purcell et al., 2007). Foram definidos como segmentos de homozigose (ROH) aqueles segmentos que apresentaram pelo menos 50 SNPs homozigotos em distâncias mínimas de 1.000 Kb por animal, permitindo um SNP heterozigoto e um SNP faltante/perdido dentro de uma janela de 50 SNPs (Scraggs et al., 2013). Posteriormente, foi feito o agrupamento dos indivíduos que compartilham segmentos de homozigose em comum no genoma.

Os segmentos de homozigose compartilhados por mais que 50% das amostras foram analisados, buscando-se identificar genes descritos para a espécie. A busca foi realizada por meio de análises nas plataformas de dados Ensembl Genome Browser (<http://www.ensembl.org/index.html>), National Center for Biotechnology Information Gene (<http://www.ncbi.nlm.nih.gov/gene>) e PANTHER (<http://www.pantherdb.org/>), para estudos das funções biológicas e vias metabólicas.

O coeficiente de endogamia (FROH) foi estimado usando a seguinte equação:

$$FROH = \frac{\sum_k \text{Tamanho}(ROH_k)}{L}$$

em que, “k” é número de ROH identificados para cada indivíduo multiplicado pelo tamanho médio de seus ROHs e “L” é o tamanho total do genoma ovino. Logo, FROH é definido como o tamanho do genoma oriundo da análise de ROH dividido pelo tamanho total do genoma.

O coeficiente de endogamia (FHOM) foi calculado com base no excesso de homozigose, como proposto por Wright (1922) como segue:

$$FHOM = \frac{O(\neq Hom) - E(\neq Hom)}{[1 - E(\neq Hom)]} = 1 - \frac{x_i(2 - x_i)}{h_i}$$

em que, $O(\neq Hom)$ e $E(\neq Hom)$ são as homozigose observada e esperada na amostra, respectivamente, e X_i é o número de cópias do alelo de referência (alelo cujo genótipo homozigoto foi codificado como "0") para o i -ésimo SNP (Yang, et al., 2012). Isso equivale a estimar a relação de um indivíduo com ele mesmo (diagonal do GRM derivado de SNP), e $h_i = 2p_i(1 - p_i)$, em que p_i é a fração observada do primeiro alelo no locus i .

O coeficiente de endogamia FGRM foi obtido pela diagonal de uma matriz de relação genômica (VanRaden 2008; VanRaden et al., 2011), calculado como:

$$G = \frac{ZZ'}{2 \sum_{l=1}^n p_l(1 - p_l)}$$

em que, Z é uma matriz de genótipo centrada e p_l é a frequência do alelo de referência no locus l . A matriz Z é obtida subtraindo-se a matriz do genótipo M (com escores genotípicos codificados como 0, 1 ou 2 para o alelo alternativo Homozigoto, heterozigoto e alelo homozigoto de referência, respectivamente). Matriz p_l , cujos elementos da coluna l são equivalentes a $2p_l$. Os elementos diagonais G ($G_{i,i}$) representam a relação de um animal consigo mesmo, incluindo assim as informações de autozigose. Segundo VanRaden et al. (2011), $G_{i,i}$ pode fornecer uma aproximação mais adequada para o coeficiente de endogamia com base no *pedigree*, quando assumido = 0,5, em vez de usar estimativas de frequências de alelos de população base (o que pode ser difícil de estimar, especialmente na ausência de dados de *pedigree*). Assim, a matriz G foi calculada utilizando frequências alélicas fixadas em 0,5.

Análises de pedigree

Para análise de *pedigree* foram utilizadas informações de registro genealógico de 428 animais da raça Santa Inês, referentes aos estados do Piauí e Maranhão. Com base nos registros, foram calculados coeficiente de endogamia (FPED) e coeficiente de parentesco médio (AR).

Coeficiente de endogamia (FPED) foi calculado utilizando-se o algoritmo proposto por Meuwissen & Luo (1992), com base na decomposição da matriz A (de relação genética aditiva), descrita por Henderson (1976): $A=LDL'$, em que L é uma matriz triangular contendo a fração dos genes de animais que derivam de seus antepassados, D é uma matriz diagonal contendo as variâncias genéticas aditivas dentro de família e L' é a transposta da matriz L. A partir da decomposição proposta por Quaas (1976), como segue:

$$A_{ii} = \sum_{j=1}^i L_{ij}^2 D_{jj}$$

em que A_{ii} é o i -ésimo elemento diagonal de A, que é igual ao coeficiente de consanguinidade do animal $i + 1$.

Coeficiente de parentesco médio (AR) é definido como a dupla probabilidade de dois alelos tomados ao acaso, um para o animal e outro para a população no *pedigree* (incluindo o do animal), serem idênticos por descendência (Goyache *et al.*, 2003). O AR de cada indivíduo no *pedigree* foi calculado como a média dos coeficientes na linha correspondente ao indivíduo em relação à matriz de parentesco A.

As matrizes G e A, e os o FGRM e FPED foram obtidos através do pacote *pedigree* do programa estatístico R.

Resultados e Discussão

O Desequilíbrio de Ligação (LD) representado pelo valor de r^2 para todos os SNPs adjacentes que estão a uma distância menor que 100 Kb foi de 0,44, superior ao encontrado por Garcia-Gómez et al. (2012), em ovinos nativos da raça Churras na Espanha (r^2 de 0,32). Isso já era de se esperar, pois a raça Santa Inês representa o grupo genético ovino mais difundido no país (Silva et al., 2012), com programas específicos de melhoramento genético e seleção (Sousa & Moraes, 2000; Lôbo, 2003). Os processos de seleção provocam aumento no nível do desequilíbrio de ligação nas regiões genômicas afetadas, já que favorecem o aumento na frequência de determinado alelo no *pool* de genes da população e, assim, aumentam a correlação com outros alelos de alta frequência, levando alelos ligados a apresentarem alta prevalência (O'Brien et al., 2014).

Foi observado redução de r^2 à medida que a distância física entre os marcadores aumentou, no entanto, a partir de distâncias superiores a 1000 Kb, o r^2 médio se estabilizou em 0,35. O decaimento de r^2 com aumento das distâncias nos marcadores é esperado, uma vez que, quanto maior for essa distância, maior será a possibilidade de ocorrerem eventos de recombinação e, portanto, a quebra da ligação entre os marcadores (Figura 2). Corroborando com os encontrados por García-Gómez et al. (2012) e Huang & Ren (2013), que observaram médias de r^2 maiores para distâncias menores.

Os coeficientes de parentesco médio tradicional e genômico obtidos a partir das matrizes A e G foram de 0,02 e 0,25, respectivamente, com coeficiente de correlação de 0,72. Este valor corroborou com vários trabalhos que verificaram correlações de alta magnitude entre as estimativas de parentesco obtidos a partir de informações de *pedigree*, pela matriz A, com os obtidos com base em dados genômicos da Matriz G (Rolf et al., 2010; Lopes et al., 2013), com valores de 0,87 e 0,78, respectivamente. Isso pode ser atribuído ao fato de que o parentesco tradicional foi obtido pela matriz A, com informações de todos os animais.

Segundo Lopes et al. (2013), quando a população inteira é analisada, o parentesco entre muitos animais não relacionados é estimado, desta forma, tanto o parentesco genômico, quanto o parentesco tradicional são próximos de zero, o que resulta em alta correlação.

Quando comparadas estimativas de parentesco tradicional, obtidas pela matriz A, com parentesco genômico, obtidos pela matriz G, foram observadas discrepâncias nas estimativas para diferentes situações, com grupos de animais com parentesco zero na matriz A e com valores variando de 0,16 a 0,8, em G, grupos de animais que apresentaram parentesco de 0,5 em A, enquanto em G foi de 0,25, e grupos de animais endogâmicos, com parentesco superior a 1,0, em G, e não endogâmicos, em A (Figura 3). Estas discrepâncias observadas são indicações claras de falhas no *pedigree* ocasionadas, provavelmente, devido às estimativas de parentesco tradicional negligenciarem os efeitos da endogamia acumulada ao longo das gerações e não considerarem os erros de *pedigree* acumulados entre as gerações passadas e de genealogia complexa com laços potencialmente desconhecidos.

Foram observadas situações em que as estimativas de parentesco genômico e tradicional se aproximaram, entretanto, o parentesco genômico apresentou grandes variações em torno das médias (Figura 3), o que é esperado devido à segregação dos genes. No entanto, se tirarmos como exemplo o parentesco de irmãos completo, que na metodologia tradicional é de 0,5, este valor pode variar quando estimado com dados genômicos, devido ao fato de que irmãos completos podem compartilhar zero IBD em cada *locus*. Se dois irmãos completos compartilham dois alelos IBD para todos os *loci*, o parentesco entre eles será 0,50. Mas, se não compartilham nenhum alelo IBD, o parentesco será zero. Portanto, é esperada uma variação em torno da média.

Variações na relação de parentesco foram observadas por Lopes et al. (2013), para parentesco de irmãos completos e meios-irmãos, de 0,08 a 0,34, com médias de 0,25 e 0,02 a 0,28, com média de 0,12, respectivamente. Segundo Hill & Weir (2011), mesmo quando o

pedigree tem perfeito relacionamento entre indivíduos de uma população, o valor previsto de parentesco pode variar devido à segregação e recombinação dos genes.

Os coeficientes de endogamia genômicos obtidos a partir dos segmentos de homozigose (FROH), pela diferença entre o número observado e o esperado de genótipos homozigóticos (FHOM) e por meio da matriz de parentescos genômica (FGRM) foram de 0,035 para FROH, apresentando valor mínimo de 0,006 e máximo de 0,203, 0,025 para FHOM apresentando valor mínimo de 0,000 e máximo de 0,27, e 0,551 para FGRM, com máximo de 0,832 e mínimo de 0,450, enquanto que para análise com base no *pedigree* FPED foi de 0,0007, com máximo de 0,25 e mínimo de 0,000 (Figura 4).

O coeficiente de endogamia estimado com base em dados de *pedigree* foi subestimado em relação aos medidos com dados genômicos, FROH, FHOM e FGRM. Uma explicação para isso é que o FPED depende da integridade e exatidão dos registros genealógicos disponíveis e, muitas vezes, a genealogia pode ser complexa e ter laços potencialmente desconhecidos. Além disso, o efeito da variação da recombinação gênica não é considerado quando a endogamia é calculada a partir de dados de *pedigree*, o que pode ser confirmando pela Figura 3, na qual se pode ser observado que o parentesco tradicional foi subestimado para a maioria dos animais, quando comparado com o parentesco genômico.

O coeficiente de endogamia (FPED) de um indivíduo depende de quanto sua ascendência é conhecida, de modo que, quanto maior for esse conhecimento, mais confiável será seu coeficiente de endogamia estimado. Segundo Mastrangelo et al. (2016), além de não levarem em consideração o efeito da variação da recombinação gênica, as estimativas com base nos registros do *pedigree* negligenciam os efeitos da endogamia acumulada ao longo das gerações e não consideram os erros de *pedigree* acumulados entre as gerações passadas.

A diferença observada entre as medidas de endogamia genômica, FHOM, FGRM e FROH, pode ser atribuída ao fato de que o cálculo de FHOM e FGRM não diferiram alelos

idênticos por descendência (IBD) de alelos idênticos por estado (IBS), os quais são ambos incluídos no cálculo da endogamia, enquanto o cálculo de FROH é com base apenas nos alelos idênticos por descendência. Além disso, FROH representa a porcentagem real do genoma que é homozigota com maior precisão, pois captura homozigose decorrente de ancestrais comuns muito distantes (por exemplo, mais de 50 gerações anteriores). FROH pode ser estimado em qualquer indivíduo genotipado, mesmo que registros de *pedigree* não estejam disponíveis. FROH também oferece a possibilidade de examinar a distribuição de homozigose em todo o genoma e identificar locais específicos no genoma que apresentam maiores níveis de homozigose (Purfield et al., 2012).

As correlações entre os coeficientes de endogamia variaram de negativas e próximas de zero a positivas de baixa magnitude (Tabela 1). Constata-se, portanto, concordância com vários trabalhos que verificaram correlações de baixa magnitude entre as estimativas do coeficiente de endogamia obtidas a partir do *pedigree* e com base em dados genômicos, (Lopes et al., 2013; Scraggs et al., 2013).

Um total de 4022 ROH foram detectados em um limiar de 50 SNPs, com tamanho médio de 6,24 Mb, com média de 14 ROH por animal (Figura 5). O tamanho médio do ROH foi maior que os observados na literatura (McQuillan et al., 2008; Kim et al., 2013). No entanto, a média do número de ROH por animal foi bem inferior.

Segmentos de homozigose com tamanho superior a 5Mb podem ser considerados como grandes, segundo Bosse et al. (2012). Tendo em vista isso, o elevado número de segmentos em homozigose encontrado com tamanho superior a 5Mb está associado ao histórico de desenvolvimento da raça, que, segundo Teixeira Neto et al. (2013), se desenvolveu a partir de estreita base genética, com 10 fundadores responsáveis por 30,51% da variabilidade genética inicial da raça Santa Inês, que é considerada por Silva et al. (2012) como o grupo genético ovino mais difundido no país.

Outro fator importante, que pode ter contribuído para a formação dos grandes segmentos em homozigose na raça Santa Inês, é o uso excessivo de alguns animais tidos como melhores geneticamente e de animais vencedores de exposições ou parentes próximos, o que reduz o tamanho efetivo, contribui para a perda da diversidade genética, aumento no coeficiente de endogamia da população e formação dos segmentos de homozigose.

Os longos segmentos de homozigose podem indicar endogamia recente na população estudada. Como os eventos de recombinação interrompem segmentos cromossômicos longos, ROH longos surgem com acúmulo de endogamia recente (até cinco gerações atrás), enquanto ROH mais curtos podem indicar efeito ancestral mais distante (até 50 gerações passadas), tais como efeito fundador da raça (Howrigan et al., 2011). Segundo Kim et al. (2013), os longos ROH podem ser resultados de gargalos de formação da raça juntamente com seleção constante e endogamia.

O elevado número de segmentos em homozigose com tamanho superior a 5Mb encontrados são indicativos de ausência de cruzamentos recente na raça Santa Inês, uma vez que o cruzamento entre raças distintas teria provocado a quebra destes segmentos (Figura 5). Isso pode ser justificado por se tratar de animais registrados junto a Associação Brasileira de Criadores de Ovinos (ARCO), a qual exige para emissão do registro a comprovação de que o animal seja filho de pais puros.

O indicativo de ausência de cruzamento recente na raça Santa Inês diverge das proposições encontradas na literatura, que afirmam que o novo Santa Inês é fruto de cruzamentos com raças exóticas como a Somalis, Sullfok, e Dorper (Paiva *et al.*, 2005; Carneiro *et al.*, 2010; Biagiotti et al., 2015).

Entre todos os segmentos de homozigose observados, não foram encontrados segmentos idênticos que fossem compartilhados por todos os animais da população em estudo. Os cromossomos 1, 2 e 3 foram os que apresentaram maior número de segmentos de

homozigose compartilhados na população, com destaque para o cromossomo 2, que apresentou mais de 600 ROH compartilhados, no entanto, os cromossomos 24 e 26 não apresentaram nenhum ROH compartilhado (Figura 6).

Os cromossomos 1, 2 e 3 apresentaram maior número de segmentos de homozigose compartilhados nos rebanhos amostrados, o que pode ser indicativo de que estes cromossomos foram favorecidos pela seleção, em que combinações específicas de alelos podem ser conservadas ao longo de muitas gerações mais frequentemente do que o esperado por acaso, conduzindo ao desenvolvimento dos segmentos em homozigose. Por outro lado, isso pode ter ocorrido simplesmente por estes serem os maiores cromossomos do genoma ovino, com maior número de marcadores, como descrito por Joaquim (2016), que em um estudo com suínos observou alto número de segmentos de homozigose compartilhados no cromossomo 1 e justificou ao fato de que este é o maior cromossomo do genoma suíno e com maior número de marcadores.

Foram identificados três segmentos de homozigose compartilhados por mais de 50% dos animais dos rebanhos amostrados (Tabela 2). Regiões de homozigose compartilhada por mais de 50% da população amostrada pode indicar que esta seja uma região conservada do genoma ovino favorecida pela seleção ou originada com a própria formação da raça.

Segundo Nothnagel et al. (2010), regiões de homozigose compartilhadas por mais de 50% dos indivíduos de uma população podem indicar a ocorrência de seleção de alta intensidade na população. Em geral, esses segmentos são resultado de uma série de mecanismos que incluem alterações cromossômicas e homozigose como consequência da endogamia e seleção natural (Yang et al., 2012).

Conclusões

Segmentos em homozigose compartilhados por mais de 50% da população indicam a ocorrência de seleção intensa para características cuja a expressão estão reguladas por genes localizados no cromossomo 16.

A endogamia calculada com base em dados de *pedigree* foi subestimada em relação a endogamia calculada com base na matriz G.

Foram verificadas falhas no *pedigree*, o que resulta em estimativas errôneas das relações de parentesco e compromete a predição de valores genéticos dos animais, quando empregados métodos tradicionais de avaliação genética.

Referências

- BIAGIOTTI, D., SARMENTO, J.L.R.2; REGO NETO, A.A. et al. Diferenciação de populações ovinas encontradas no estado do Piauí Diferenciação de populações ovinas encontradas no estado do Piauí. **Archivos de zootecnia** vol. 64,n. 245, p. 16. 2015.
- BOSSE M, MEGENS HJ, MADSEN O, PAUDEL Y, FRANTZ LA, SCHOOK LB, et al. Regions of homozygosity in the porcine genome: consequence of demography and the recombination landscape. **PLoS Genet.** 2012;8(11):e1003100.
- CARNEIRO, H., LOUVANDINI, H., PAIVA, S.R., MACEDO, F., MERNIES, B., MCMANUS, C. Morphological characterization of sheep breeds in Brazil, Uruguay and Colombia. **Small Ruminant Research**, v.94, p.58-65, 2010.
- FORNI, S., AGUILAR, I., MISZTAL, I. Different genomic relationship matrices for singlestep analysis using phenotypic, *pedigree* and genomic information. **Genetics**,
- GARCÍA-GÁMEZ, E., SAHANA, G., GUTIÉRREZ-GIL, B AND ARRANZ, J.J. Linkage disequilibrium and inbreeding estimation in Spanish Churra sheep. **BMC Genetics** 13, 43. 2012.
- GOYACHE, F. et al. Using *pedigree* information to monitor genetic variability of endangered populations: the Xalda sheep breed of Asturias as an example. **Journal of Animal Breeding and Genetics**, v.120: p.95-103, 2003.
- HENDERSON, C.R. A simple method for computing the inverse of a numerator relationship matrix used in the prediction of breeding values. **Biometrics** 32, 69-83 IMSL (1984) **International Mathematical and Statistical Libraries**. Houston, TX, ed. 9.2. 1976.
- HILL WG, WEIR BS. Variation in actual relationship as a consequence of Mendelian sampling and linkage. **Genetics Research**, 93, 47–64, 2011.
- HOWRIGAN, DP, SIMONSON MA AND KELLER MC. Detecting autozygosity through runs of homozygosity: a comparison of three autozygosity detection algorithms. **BMC Genomics** 12, 460 2011.

JOAQUIM, L.B. **Estrutura genômica de uma população de suínos base Landrace.** 37p. Dissertação (Zootecnia). UNESP- Universidade Estadual Paulista, Jaboticabal. 2016.

KIM ES, COLE JB, HUSON H, WIGGANS GR, VAN TASSELL CP, Crooker BA, Liu G, Da Y and Sonstegard TS. Effect of artificial selection on runs of homozygosity in U.S. Holstein cattle. **PLoS One** 8, 2013.

LOPES, M. S.; SILVA, F. F.; HARLIZIUS, B.; DUIJVESTIEN, N.; LOPES, P. S. GUIMARÃES, S. E. F.; KNOL, E. F. Improved estimation of inbreeding and kinship in pigs using optimized SNP panels. **BMC Genetics**, v. 14, 2013.

MASTRANGELO, S., TOLONE, M., GERLANDO R. et al. Genomic inbreeding estimation in small populations: evaluation of runs of homozygosity in three local dairy cattle breeds. **Animal**, 1(5):1-9, 2016.

MCQUILLAN, R., LEUTENEGGER, A., ABDEL-RAHMAN, R. et al. Runso homozygosity in European populations. **Am.J. Hum. Genet.** 83, 359–372, 2008.

MEUWISSEN, T.H.E; LUO, Z. Computing inbreeding coefficients in large populations. **Genetics Selection Evolution**, v.24, p.305-313, 1992

NOTHNAGEL M, LU TT, KAYSER M, KRAWCZAK M. Genomic and geographic distribution of SNP-defined runs of homozygosity in Europeans. **Hum Mol Genet** 19(15): 2927-2935. doi:10.1093/hmg/ddq198. PubMed: 20462934, 2010.

O'BRIEN, A. M. P.; MÉSZÁROS, G.; UTSUNOMIYA, Y. T.; SONSTERGARD, T. S.; GARCIA, F. J.; TESSELL, C. P. V.; CAVALHEIRO, R.; SILVA, M. V. B.; SOLKNER, J. Linkage disequilibrium levels in *Bos indicus* and *Bos taurus* cattle using medium and high density SNP chip data and different minor allele frequency distributions. **Livestock Science**, v. 166, p. 121-132, 2014.

PAIVA, S. **Caracterização da diversidade genética de ovinos no Brasil com quatro técnicas molecular.** 2005. 118 f. Tese (Programa de Pós-Graduação em Genética e melhoramento), Universidade Federal de Viçosa, Viçosa, 2005.

PURCELL, S., NEALE, B., TODD-BROWN, K., THOMAS, L., FERREIRA, M.A., BENDER, D. PLINK: a tool set for whole-genome association and population-based linkage analyses. **Am J Hum Genet** 81. 2007.

PURFIELD DC, BERRY DP, MCPARLAND S AND BRADLEY DG. Runs of homozygosity and population history in cattle. **BMC Genetics** 13, 70, 2012.

QUAAS, R.L. Computing the diagonal elements and inverse of a large numerator relationship matrix. **Biometrics**, v.32, p.949-953, 1976.

ROLF, M.M; TAYLOR, J.F.; SCHNABEL, R.D.; MCKAY, S.D.; MCCLURE, M.C.; NORTH CUTT, S.L.; KERLEY, M.S.; WEABER, R.L.: Impact of reduced marker set estimation of genomic relationship matrices on genomic selection for feed efficiency in Angus cattle. **BMC genetics** 2010, 11(1):24.

- SCRAGGS, E.; ZANELLA, R.; WOJTOWICZ, A.; TAYLOR, J.F.; GASKINS, C.T.; REEVES, J.J.; de AVILA, J.M.; NEIBERGS, H.L. Estimation of inbreeding and effective population size of full-blood wagyu cattle registered with the American Wagyu Cattle Association. **Journal of Animal Breeding and Genetics**, v. 131, n. 1, p.3-10, 2013.
- SILVA, L. S. A. da; FRAGA, A. B., SILVA, F. de L. da; BEELEN, P. M. G.; SILVA, R. M. de O. Growth curve in Santa Inês sheep. **Small Ruminant Research**, v. 105, p. 182-185, 2012.
- TEIXEIRA NETO, R. M. et al. Parâmetros populacionais da raça ovina Santa Inês no Brasil. **Pesq. agropec. bras.** v. 48. p. 1589 a 1595, 2013.
- VANRADEN, P.M. Efficient methods to compute genomic predictions. **J Dairy Sci.** 91: 4414±4423.2008.
- VANRADEN PM, OLSON KM, WIGGANS GR, COLE JB, TOOKER ME: Genomic inbreeding and relationships among Holsteins, Jerseys, and Brown Swiss. **J Dairy Sci**, 94(11):5673–5682. 2011.
- WANG, L.; SORENSEN, P.; JANSS, L.; OSTERSEN, T.; EDWARDS, D. Genomewide and local pattern of linkage disequilibrium and persistence of phase for 3 Danish pig breeds. **BMC Genetics**, v. 14, p. 115-126, 2013.
- WRIGHT, S., 1922 WRIGHT, S., 1943 WRIGHT, S., 1951 Coefficients of inbreeding and relationship. **Am. Nat.** 56: 330-338.
- YANG, H. C.; CHANG, L. C.; LIANG, Y. J.; LIN, C. H.; WANG, P. L. A Genome-wide Homozygosity Association Study Identifies Runs of Homozygosity Associated with Rheumatoid Arthritis in the Human Major Histocompatibility Complex. **Plos One**, v. 7, 34840, 2012.

Tabela 1. Correlação de *Pearson* entre coeficientes de endogamia genômicos e tradicional em ovinos da raça santa Inês

| | FHOM | FROH | FRGM |
|------|-------|-------|-------|
| FROH | -0,08 | | |
| FRGM | 0,03 | -0,15 | |
| FPED | 0,32 | -0,02 | -0,06 |

FHOM = coeficiente de endogamia com base na diferença entre o número observado e o esperado de genótipos homozigóticos; FROH = coeficiente de endogamia com base nos seguimentos de Homozigose; FRGM = coeficiente de endogamia baseado na matriz de relacinamito genômica; FPED= coeficiente de endogamia baseado na matriz de relacinamento tradicional de *pedigree*.

Tabela 2. Número de animais e tamanho dos segmentos de homozigose compartilhados por mais de 50% da população por cromossomos da raça Santa Inês

| Segmento | CHR | N | % | Início | Fim | Tamanho |
|----------|-----|-----|------|----------|----------|---------|
| 1 | 16 | 148 | 54,6 | 41681642 | 41711634 | 29993 |
| 2 | 16 | 146 | 53,9 | 41737441 | 41804913 | 67473 |
| 3 | 16 | 146 | 53,9 | 42806888 | 42894008 | 87121 |

CHR=Cromossomo; N=Número de animais com o mesmo segmento; %=Porcentagem de animais com o mesmo segmento; Tamanho do segmento compartilhado (bp); Início do segmento compartilhado (bp); Final do segmento compartilhado

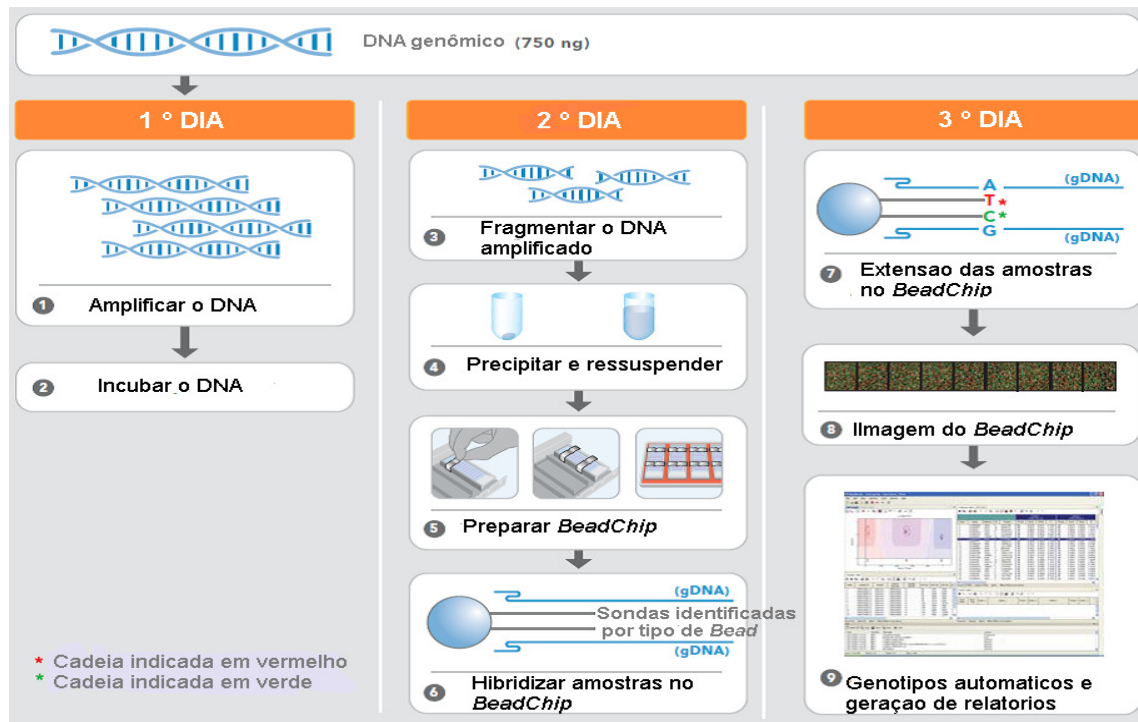


Figura 1- Genotipagem com tecnologia *Infinium* da *Illumina*

Fonte: (<http://www.illumina.com>)

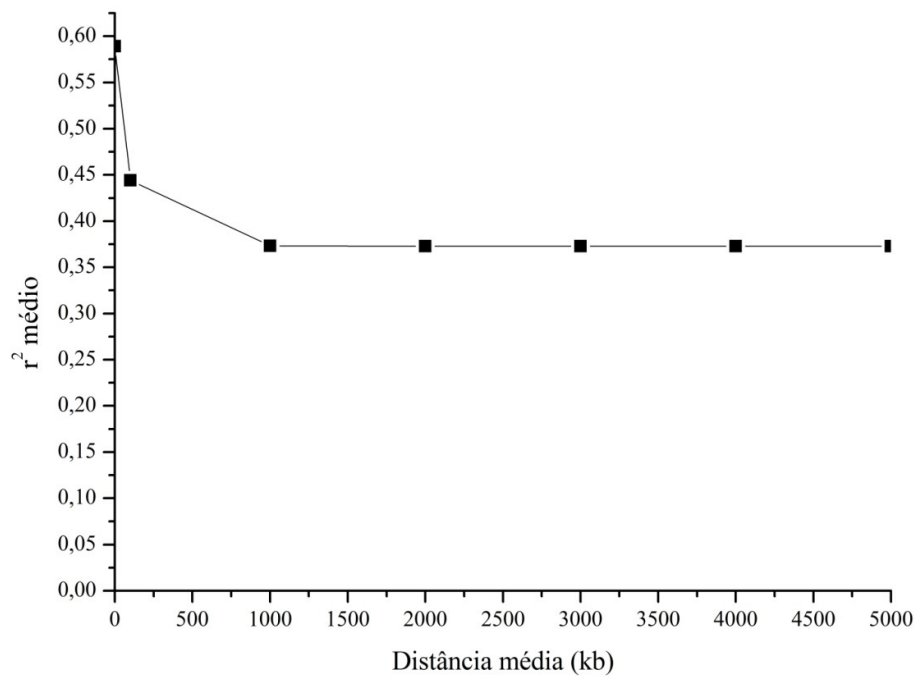


Figura 2. Média de desequilíbrio de ligação (r^2) ao longo de diferentes distâncias (kb) entre marcadores em ovinos da raça Santa Inês.

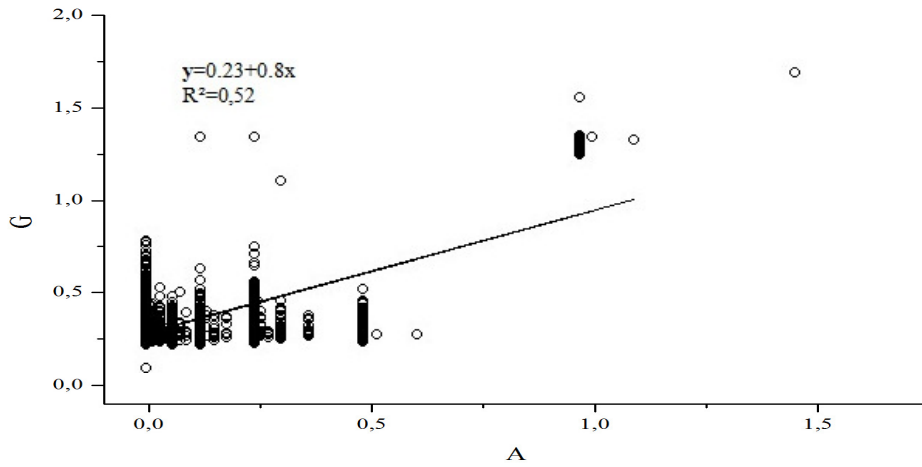


Figura 3. Regressão entre parentesco genômico obtido pela matriz (G) com o parentesco tradicional com base no *pedigree* obtido pela matriz (A) em animais da raça Santa Inês.

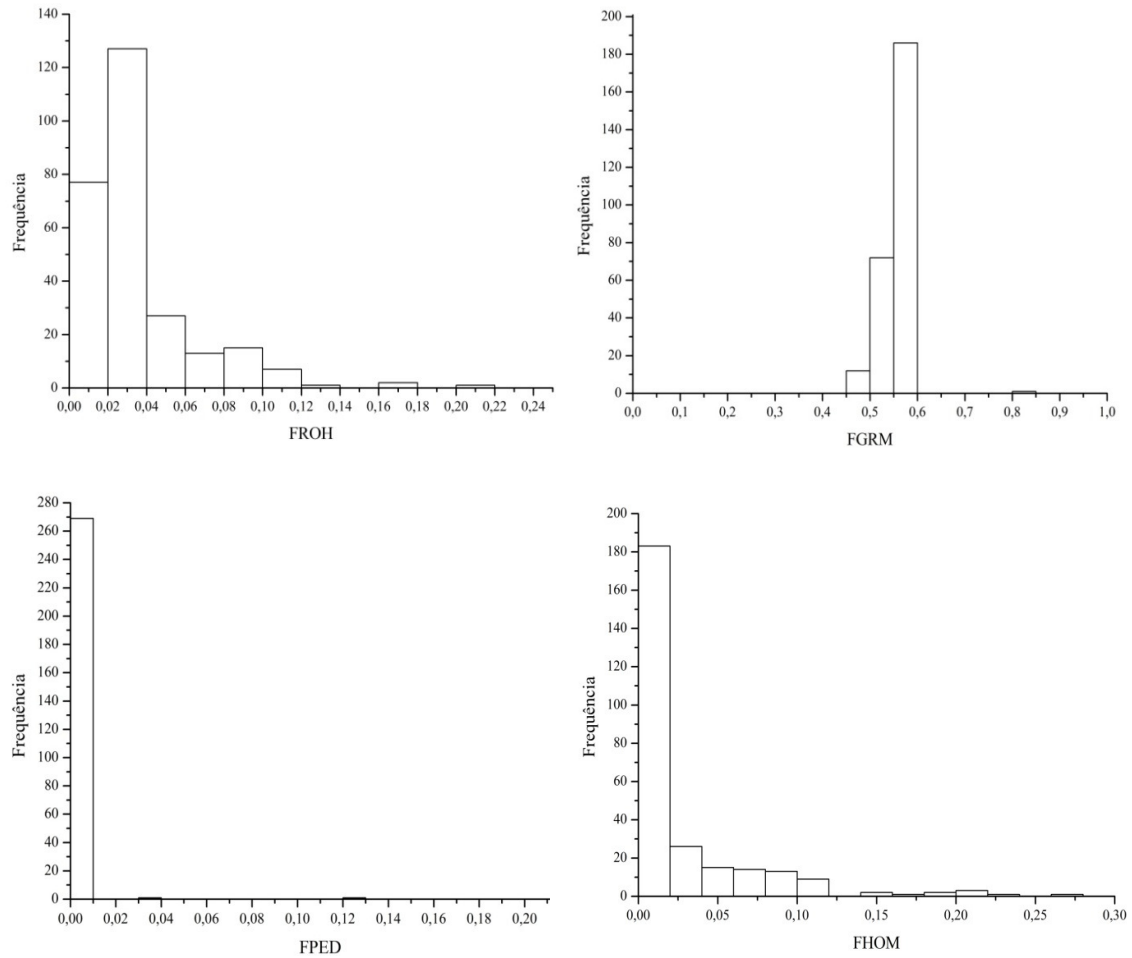


Figura 4. Histogramas do coeficiente de endogamia de ovinos da raça Santa Inês utilizando dados genômicos a partir dos segmentos de homozigose (FROH), homozigose esperada e observada (FHOM), matriz de parentesco genômica (FGRM) e registros de *pedigree* (FPED).

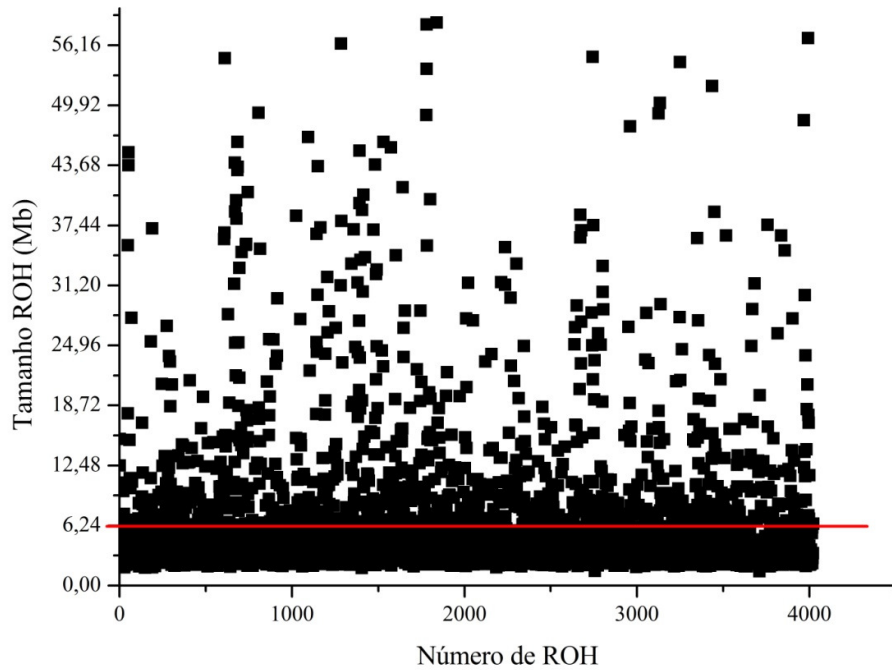


Figura 5. Relação entre o número e tamanho médio dos segmentos de homozigose ROH, com média representada pela linha vermelha em ovinos da Raça Santa Inês.

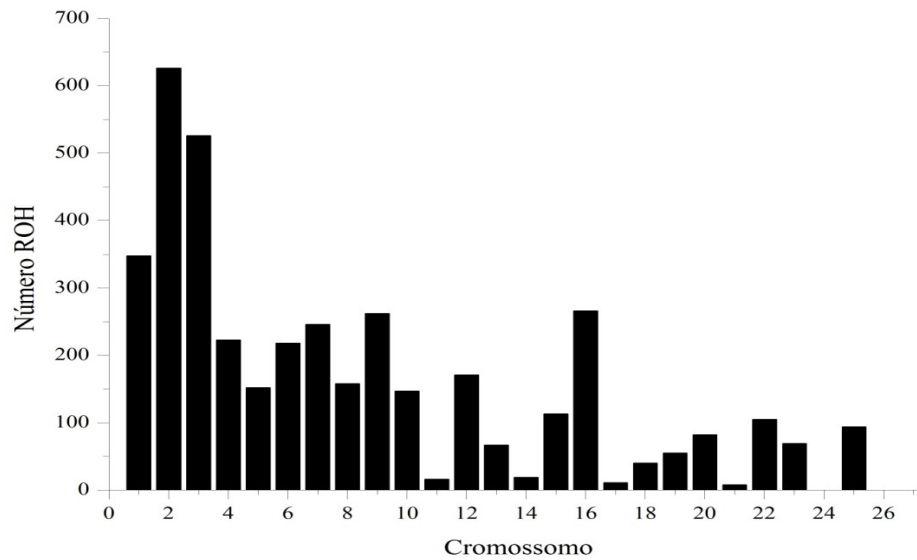


Figura 6. Número de regiões de homozigose compartilhadas entre indivíduos por cromossomos da raça Santa Inês.

4 CAPÍTULO II

ASSOCIAÇÃO GENÔMICA AMPLA PARA CARACTERÍSTICAS DE TAMANHO CORPORAL EM OVINOS DA RAÇA SANTA INÊS

Associação genômica ampla para características de tamanho corporal em ovinos da raça Santa Inês

Aurino de Araujo Rego Neto⁽¹⁾ e José Lindenberg Rocha Sarmiento⁽²⁾

⁽¹⁾Pós Graduando do Programa de Pós-Graduação em Ciência Animal – UFPI. e-mail:

aurinorego@hotmail.com

⁽²⁾Professor da UFPI/DZO e Bolsista do CNPq, e-mail: sarmiento@ufpi.edu.br

Resumo: Estudos de associação genômica ampla (GWAS) possibilitam identificar regiões de identidade no genoma associadas a fenótipos e têm sido aplicados em muitas características de interesse econômico em diferentes espécies. Objetivou-se com esta pesquisa identificar regiões genômicas associadas a características de tamanho corporal de animais da raça Santa Inês. Para tanto, foram utilizados informações de 428 animais criados nos estados do Piauí e Maranhão com registro na Associação Brasileira de Criadores de Ovinos, dos quais foram genotipados 271, utilizando-se SNPs Chip de alta densidade Ovine 50K da Illumina. Após o controle de qualidade, foram utilizados 43.164 SNPs. Para verificar a associação marcadores SNPs com características de tamanho corporal realizou-se análise de associação genômica ampla (GWAS) por meio da metodologia GWAS passo único (ssGWAS) para estimar os efeitos de marcadores e associá-los às características peso a idade adulta (PA), comprimento corporal (CC), altura da cernelha (AC), circunferência torácica (CT), comprimento da perna (CP) e perímetro da perna (PP). Observaram-se associações com PA, CP e PP no cromossomo 6, com CT nos cromossomos 4, 7 e 13 e com AC e CC no cromossomo 4. As regiões identificadas neste estudo apresentam vários genes (BMP3, FGF5, SPP1, ABCG2, ALVA, BMPR1B, CCKAR, HGF, CACNA2D1) com conhecimento biológico descrito que poderão

auxiliar na melhor compreensão da expressão destas características, assim como, poderão auxiliar nas tomadas de decisões em programas de seleção da raça Santa Inês.

Termos para indexação: perímetro da perna, peso, regiões genômica, SNP, ssGWAS

Genome-wide association for body size traits in Santa Inês sheep

Abstract. Genome-wide association studies (GWAS) allows for the identification of regions of identity throughout the genome associated with phenotypes. Thereby, GWAS have been applied to several economically important traits in different species. This study aimed to identify genomic regions associated with body size traits in Santa Inês sheep. Therefore, information of 428 animals raised in the states of Piauí and Maranhão, and registered with the Brazilian Association of Sheep Breeders, was used. From those animals, 271 were genotyped using the high density Illumina Ovine SNP50K BeadChip. After quality control, 43,164 SNPs were used for genomic analysis. A genome-wide association (GWA) analysis was performed using the single-step GWAS (ssGWAS) method, in order to estimate markers effects and associate them to adult body weight (PA), body length (CC), height at withers (AC), thoracic circumference (CT), leg length (CP), and leg perimeter (PP). Associations with PA, CP, and PP (on chromosome 6), with CT (on chromosomes 4, 7 and 13), and with AC and CC (on chromosome 4) were observed. The regions identified in this study showed several genes (BMP3, FGF5, SPP1, ABCG2, ALVA, BMPR1B, CCKAR, HGF, CACNA2D1) biologically described, which would help in the better understanding on the expression of the traits in study, and would help in decision-making in Santa Inês selection programs.

Index terms: genomic regions, leg perimeter, SNP, ssGBLUP, weight.

Introdução

As características de tamanho corporal estão diretamente relacionadas com crescimento e produção de carne em ovinos. São características que apresentam estimativa de herdabilidade variando de moderada a alta magnitude, indicando que o processo de seleção para tais características pode ser feito de forma eficiente. O tamanho do corpo se encontra sob seleção positiva na maioria dos estudos publicados (Kingsolver & Pfennig, 2004; Kingsolver & Diamond, 2011).

A relação entre o tamanho corporal e produção de carne é tradicionalmente realizada pela metodologia BLUP, que tem como base informações de *pedigree* para estimar a relação entre os indivíduos (Morrissey et al., 2012). No entanto, esta metodologia é limitada, pois não consegue identificar o número nem as posições físicas das regiões genômicas que contribuem para a variação, nem pode revelar como elas afetam as características (Visscher et al., 2007).

Os avanços tecnológicos aliados ao fator custo-benefício possibilitaram a utilização de painéis densos de polimorfismos de nucleotídeo único (“*Single Nucleotide Polymorphisms*”), denominados SNPs, para os estudos de associação genômica ampla, o que possibilita sua aplicação para auxiliar no melhoramento genético de espécies domésticas (Zhang et al., 2012). Desta forma, a utilização de informação genômica tornou-se um procedimento padrão para a avaliação animal nos programas de melhoramento genético.

Estudos de associação genômica ampla (GWAS) têm sido aplicados a muitas características de interesse econômico, em diferentes espécies, visando identificar regiões de identidade no genoma associadas a fenótipos. Tradicionalmente, as análises de GWAS são realizadas por metodologias que testam um marcador por vez como efeito fixo, assumindo ausência de estrutura da população, que é muito improvável de acontecer quando se está trabalhando com dados reais. Além disso, como o número de marcadores a serem testados é muito grande, os vários testes realizados podem incorrer em erros, o que requer o emprego de

métodos de ajustes para esses múltiplos testes como a correção de Bonferroni (Silva et al.,2017).

Uma alternativa ao GWAS tradicional consiste em integrar todos os genótipos, fenótipos e informações de *pedigree* em uma única etapa permitindo, assim, o uso de diferentes fontes de informações e todas as relações de parentesco simultaneamente. Com esta abordagem, todos os SNPs são considerados ao mesmo tempo, junto com todos os fenótipos dos animais genotipados e não genotipados, pelo método de passo único (ssGWAS) (Wang et al., 2012).

Frente a esta situação, objetivou-se com esta pesquisa identificar associações entre regiões genômicas com características de tamanho corporal, de peso a idade adulta, altura da cernelha (AC), comprimento corporal (CC), circunferência torácica (CT), comprimento da perna (CP) e perímetro da perna (PP), em ovinos da raça Santa Inês.

Material e Métodos

Foi utilizado um conjunto de dados com informações de *pedigree* de 428 animais da raça Santa Inês, pertencentes a 6 rebanhos localizados na região Meio-Norte do Brasil, composta pelos estados do Piauí e Maranhão. Dos animais em estudo, 152 tinham registros de peso à idade adulta (PA), 267 comprimento corporal (CC), 256 altura da cernelha (AC), 267 circunferência torácica (CT), 267 comprimento da perna (CP) e 267 perímetro da perna (PP).

A extração do DNA foi realizada com o Kit *Qiagen DNeasy Blood & Tissue*, segundo o protocolo estabelecido pelo fabricante. Terminada a extração, a qualidade do DNA foi verificada em gel de agarose a 1% em tampão SB 1X por 60 minutos a 60V. Os géis foram posteriormente visualizados e registrados em fotodocumentador (*L-PIX Loccus Biotecnologia*). O DNA obtido foi estocado em freezer a -20°C até as análises subsequentes.

Após a extração, foi realizada a quantificação do DNA, para isso foi utilizado o Fluorespectofotômetro NanoDrop 3300 com o kit *Intercalante AccueBlue dsDNA*, que estima a concentração de DNA das amostras, essencial para preparação das amostras para genotipagem.

A genotipagem foi realizada para 271 animais utilizando o chip de alta densidade *BeadChip OvineSNP50k*, contendo 54.241 SNPs uniformemente espaçados nos cromossomos. O chip de SNPs foi fabricado de acordo com a tecnologia Infinium da Illumina™, usando a plataforma iScan. O protocolo de genotipagem foi seguido de acordo com que foi estabelecido pelo fabricante (disponível em www.illumina.com).

O controle de qualidade dos genótipos foi realizado com o pacote HapEstXXR do programa estatístico R, de modo que, considerou-se SNPs localizados em cromossomos autossômicos e com “Genecall Score” maior que 0,70, *Minor Allele Frequency* (MAF) < 0,05, além de *Call Rate* < 0,95 e *Equilíbrio de Hardy-Weinberg* (HWE) < 10⁻⁵.

Para análise de associação genômica ampla (GWAS) foi utilizado a metodologia do de passo único (ssGWAS), proposta por Wang et al. (2012), considerando modelo unicaracterística, como segue:

$$Y = Xb + Za + e,$$

em que: Y é o vetor de observações fenotípicas; X é a matriz de incidência relacionando os fenótipos aos efeitos fixos; b é o vetor de efeitos fixos que incluiu grupos contemporâneos (propriedade, ano de nascimento e tipo de nascimento), sexo e idade do animal dividida em classes (I, animais com idade maior que um ano e menor ou igual a dois anos; II, animais com idade maior que dois anos e menor ou igual a quatro anos; e III, animais com idade maior que quatro anos); Z é a matriz de incidência que relaciona os fenótipos aos animais; a é o vetor de efeitos genético aditivo dos animais; e é o vetor de efeitos residuais.

As variâncias genética e residual foram obtidas como segue:

$$\text{var} \begin{bmatrix} a \\ e \end{bmatrix} = \begin{bmatrix} H \sigma_a^2 & 0 \\ 0 & I \sigma_e^2 \end{bmatrix}$$

em que: σ_a^2 é a variância genética aditiva e σ_e^2 é a variância residual. H é a matriz que combina a matriz de *pedigree* e a matriz de informação genômica, conforme apresentado por Aguilar et al. (2010), I é uma matriz identidade e A inversa da matriz H é:

$$H^{-1} = A^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & G^{-1} - A_{22}^{-1} \end{bmatrix}$$

em que A^{-1} é a inversa da matriz de parentesco baseada na informação de *pedigree*, G^{-1} é a inversa da matriz de parentesco genômico, que foi construída e dimensionada como descrito por VanRaden (2008), e A_{22}^{-1} é a inversa da matriz de parentesco baseada no *pedigree* para animais genotipados. A matriz H foi construída dimensionando G baseada em A_{22} , considerando que a média da diagonal de G é igual à média da diagonal de A_{22} e a média dos elementos fora da diagonal de G é igual à média dos elementos fora da diagonal em A_{22} .

O efeito de animal foi decomposto em genotipados (ag) e não genotipados, como descrito por Wang et al. (2012) com o efeito de animal genotipado:

$$ag = Zu$$

em que: Z é uma matriz que relaciona genótipos de cada locus e u é um vetor de efeitos de marcador.

A variância de efeitos de animal foi assumida como:

$$\text{var}(a_g) = \text{var}(Zu) = ZDZ' \sigma_u^2 = G^* \sigma_a^2,$$

em que: D é uma matriz diagonal de pesos para variâncias de marcadores ($D = I$ para GBLUP), σ_u^2 é a variância genética aditiva capturada por cada marcador SNP quando a matriz de parentesco ponderada (G^*) foi construída sem peso.

Assim, os efeitos de SNP foram obtidos seguindo a equação descrita por Wang et al. (2012), como segue:

$$\hat{u} = \lambda DZ' G^{*-1} \hat{a}_g = ZD'[ZDZ']^{-1} \hat{a}_g$$

em que: λ foi definida por VanRaden et al. (2008) como uma constante de normalização, como descrito abaixo.

$$\lambda = \frac{\sigma_u^2}{\sigma_a^2} = \frac{1}{\sum_{i=1}^M 2p_i(1-p_i)}$$

em que: M é o número de SNPs e p_i a frequência do segundo alelo no i -ésimo SNP.

A porcentagem de variância genética explicada pela i -ésima região foi calculada, como descrito por (Wang et al., 2014).

$$\frac{\text{var}(a_i)}{\sigma_a^2} \times 100\% = \frac{\text{var}(\sum_{j=i}^{50} z_j \hat{u}_j)}{\sigma_a^2} \times 100\%$$

em que: a_i é o valor genético da i -ésima região, que consiste de 50 SNPs consecutivos, σ_a^2 é a variância genética total, Z_j é vetor de conteúdo de SNP do j -ésimo SNP para todos os indivíduos e \hat{u}_j é o efeito de marcador do j -ésimo dentro da i -ésima região.

As análises foram processadas nos programas da família BLUPF90 (Misztal et al., 2002). Os resultados das análises de GWAS foram apresentados com base na proporção de variância explicada por janela de 50 SNPs adjacentes, com média de 594Kb.

Os segmentos cromossômicos que explicaram 1% ou mais da variância genética aditiva foram selecionados para explorar e determinar possíveis *loci* de características quantitativas (QTL) ou genes já conhecidos no genoma ovino. A ferramenta visualizador de mapas do genoma ovino foi utilizado para a identificação de genes, disponível no "National Center for Biotechnology Information" (NCBI - <http://www.ncbi.nlm.nih.gov>) em: (<http://www.ncbi.nlm.nih.gov/genome/?term=sheep>).

A classificação de genes de acordo com sua função biológica e identificação de vias metabólicas foram realizadas utilizando-se o "Banco de Dados para Anotações, Integrated

Discovery (DAVID) v. 6.7° (<http://david.abcc.ncifcrf.gov/>), GeneCards (<http://Www.genecards.org/>), e UniProt (<http://uniprot.org>).

Resultados e Discussão

As herdabilidades variaram de moderada a alta magnitude (Tabela 1), o que pode ser atribuído a metodologia ssGBLUP, utilizada para estimação, uma vez que a mesma, permite utilizar informações de todos os genótipos, fenótipos e de *pedigree* em uma única etapa permitindo, assim, o uso de diferentes fontes de informações e todas as relações de parentesco simultaneamente, contribui para aumentar a acurácia da estimação dos parâmetros genéticos. Segundo Kingsolver e Diamond (2011), características de tamanho corporal apresentam estimativa de herdabilidade variando de moderada a alta magnitude.

Para as características peso a idade adulta, comprimento da perna e perímetro da perna foram identificadas 3 regiões genômicas, no cromossomo OAR6, que explicaram mais de 1% da variância genética aditiva, nas quais foram identificados 58 genes (Figuras 1 a 3), corroborando com Matika et al (2016), que observaram SNPs significativos no cromossomo OAR6 associados com peso ósseo, peso de gordura, peso vivo e densidade muscular, em ovinos da raça Scottish Blackface,

Dentre os genes encontrados na região que explicou 1,08% da variância genética aditiva para peso a idade adulta (Tabela 2), foi observado o gene BMP3, que codifica proteína morfogenética óssea três, que em bovinos induz a formação de cartilagem e osso, as quais, segundo Daluiski et al. (2001), são membros da superfamília do fator de crescimento (TGF-beta). Muitas BMPs são produzidas no osso e apresentam atividade osteogênica, sugerindo que elas podem ser determinantes da massa óssea. No entanto, não foi identificado nenhum estudo em ovinos.

O gene FGF5, que codifica a proteína de fator de crescimento fibroblástico, funciona como um inibidor do alongamento do pêlo, identificado no cromossomo 6 por Kijas et al. (2012), em análise genômica de raças de ovinos.

Para a região genômica que explicou 1,1% da variância genética para comprimento da perna, foram identificados vários genes com funções relevantes, SPP1, ABCG2, BMPR1B, ALVA, DMP1, CSN1S1, CSN2, CSN1S2 e CSN3 (Tabela 3).

O gene SPP1 codifica a fosfoproteína 1 segregada, conhecida como osteopontina, que desempenha papel na formação óssea em diferentes espécies, incluindo humanos e camundongos (Denninger et al., 2015; De Peppel & Leeuwen, 2015). O gene ABCG2 codifica uma proteína associada à resistência ao câncer de mama (BCRP) em humanos, que pertence à superfamília dos transportadores de sinal de ligação de ATP (*ABC-ATP-binding cassette*). Em bovinos, o gene ABCG2 tem sido um dos candidatos posicionais para estudos relacionados a características de produção de leite. Em bovinos de leite, Ron et al. (2001) detectaram um QTL no cromossomo 6 (BETA6) associado ao teor de proteína e de gordura do leite. Gutierrez-Gil et al. (2014) compararam cinco raças leiteiras de ovelhas com cinco raças de corte e identificaram uma assinatura de seleção no cromossomo OAR6, que engloba os genes ABCG2 e SPP1.

BMPR1B é um gene que codifica proteínas morfogenéticas ósseas (BMP), que são proteínas não colágenas, presentes na matriz óssea, responsáveis pelo recrutamento celular e indução da diferenciação das células mesenquimais não diferenciadas em condroblastos e osteoblastos, produzindo novo tecido ósseo, *in vivo* e *in vitro* (Oleskovicz et al., 2004). Outro gene relacionado ao tecido ósseo encontrado foi o DMP1, que codifica a proteína da matriz de dentina, proteína específica para os ossos e dentes inicialmente identificada a partir da dentina mineralizada, que atua como um componente transcricional para a ativação de genes específicos dos osteoblastos, tais como a osteocalcina (Eapen et al., 2010).

Outro gene relacionado com o comprimento da perna foi o ALVA, que codifica a proteína albumina, considerada a mais abundante no sangue humano. Esta proteína funciona na regulação da pressão osmótica coloidal do plasma sanguíneo e atua como uma proteína transportadora para uma vasta gama de moléculas endógenas, incluindo hormônios, ácidos graxos e metabólitos, bem como fármacos exógenos (Brown et al., 1989). Foram identificados quatro genes da caseína (CSN1S1, CSN2, CSN1S2 e CSN3) que têm efeito positivo sobre a produção de proteína do leite e são importantes para a estrutura e estabilidade das micelas de caseína (Alexander et al., 1988).

Dentre os genes encontrados na região que explicou mais de 1,61% da variância genética para perímetro da perna (Tabela 4), pode se destacar o gene CCKAR, que codifica um receptor acoplado à proteína G, que se liga a membros não sulfatados da família de colecistoquinina (CCK) de hormonas peptídicas. Este receptor é um importante mediador fisiológico da secreção da enzima pancreática e da contração do músculo liso da vesícula biliar e do estômago. No sistema nervoso central e periférico este receptor regula a saciedade, que controla a ingestão de alimentos e a liberação de beta-endorfina e dopamina (Kazmi, et al., 2016).

Vários estudos relataram SNPs em regiões do cromossomo OAR6, envolvido no crescimento e tamanho corporal, por exemplo, na composição corporal e características de qualidade da carne em frango (Liu et al., 2013), tamanho de vitelo ao nascimento e estatura adulta em bovinos (Sahana et al., 2015), altura e tamanho corporal de cavalos (Tetens et al., 2013), varredura seletiva para o crescimento de suínos (Rubin et al., 2012), e em ovinos para características musculares, de gordura e ósseas (Matika et al., 2016).

Para as características de comprimento corporal, altura da cernelha e circunferência torácica, foram identificadas regiões genômicas que explicaram mais de 1% da variância

genética aditiva no cromossomo OAR4, sendo que para circunferência torácica foram observadas regiões nos cromossomos OAR7 e OAR 13 (Figuras 4 a 6).

Foi identificada uma região no cromossomo OAR4 que explica 1,31% da variância genética aditiva para comprimento corporal, na qual contém 11 genes candidatos descritos no genoma ovino (Tabelas 5).

Foi identificada uma região genômica em comum para as características de comprimento corporal e altura da cernelha no cromossomo OAR4, dentro das regiões genômicas que explicaram mais de 1% da variância genética, localizado na posição 35,1MB a 37,8MB, na qual estão localizados 10 genes candidatos com múltiplos efeitos ou associados às duas características ao mesmo tempo (SEMA3D, LOC105607853, SEMA3A, LOC105614979, LOC105614978, SEMA3E, LOC105615152, LOC105614980, PCLO e LOC101114098).

Dentre os genes identificados na região cromossômica em comum para as características de comprimento corporal e altura da cernelha, estão 3 genes da família das semaforinas de classe 3 (SEMA3D, SEMA3A e SEMA3E), que desempenha papel importante na sinalização célula-célula e na regulação da migração de um amplo espectro de tipos de células, regula a migração de timóitos na medula, regula a migração das células endoteliais, desempenha um papel importante na garantia da especificidade da formação de sinapses, mede a sinalização anti-angiogênica em resposta à SEMA3E e é necessário para o desenvolvimento normal do coração e da vasculatura (Rimar et al., 216). A proteína codificada pelo gene PCLO faz parte da matriz pré-sináptica do citoesqueleto, que está envolvida no estabelecimento de zonas sinápticas ativas e no tráfico de vesículas sinápticas. As variações neste gene foram associadas à desordem bipolar e desordem depressiva principal (Ahmed et al., 2015).

Para altura da cernelha, foram identificadas duas regiões genômicas próximas, no cromossomo OAR4, que explicaram 1,91 e 1,15% da variância genética aditiva, nas quais estão localizados, respectivamente, 17 e 11 genes candidatos (Tabelas 6).

Dos genes contidos nas duas regiões no cromossomo 4 (Tabela 5), pode-se destacar o HGF, descrito como fator de crescimento de hepatócitos, que atua como um fator de crescimento para um amplo espectro de tecidos e tipos de células. O gene HGF expressa o fator de crescimento hepático, que é produzido e secretado pelo tecido adiposo branco e estimula a diferenciação e desenvolvimento de adipócitos (Fonseca-Alaniz et al., 2006). O gene GNAT3, que codifica a subunidade alfa da proteína de ligação a nucleótidos de Guanina (proteína G), desempenhando um papel proeminente na transdução do sabor amargo e doce. O gene CACNA2D1 (NM_001205987) que codifica um membro da família da subunidade alfa2/delta, uma proteína do complexo do canal de cálcio dependente de voltagem. Em bovinos, este gene foi mapeado em uma região genômica de sete QTLs para peso da carcaça, ganho médio diário (Casas et al., 2001), estatura e porcentagem de osso e razão carne/osso (Gutiérrez-Gil et al., 2009).

Para circunferência torácica, foram identificadas 3 regiões genômicas que explicaram 1,13, 1,07 e 1,02% da variância genética aditiva, localizadas nos cromossomos 13, 7, e 4, com respectivamente 22, 30 e 40 genes candidatos em cada região genômica (Tabela 7).

As características de altura da cernelha e circunferência torácica apresentaram um segmento em comum no cromossomo OAR4 dentro das regiões genômicas, que explicaram mais de 1% da variância genética aditiva (Tabelas 6 e 7), localizado na posição 39,8MB a 41,5MB, na qual foram identificados 11 genes candidatos (LOC105614987; SEMA3C; LOC10102230; LOC101114861; LOC105614988; LOC105614989; LOC101115115; GNAT3; LOC101115786; GNAI1; e LOC105613103). Dentre esses genes estão LOC10102230 e LOC101114861, que codificam aglicoproteína de plaquetas, conhecida

como integrina $\alpha_{IIb}\beta_3$, que é uma proteína do grupo das integrinas encontrada em plaquetas. É um receptor para o fibrinogênio e auxilia na ativação plaquetária. As plaquetas secretam uma série de substâncias, tais como: ADP, serotonina, fibrinogênio, fator de von Willebrand, fibronectina, fatores de crescimento (fatores de crescimento derivados das plaquetas, fator de crescimento alfa e fator de crescimento beta), pró-coagulantes (fator plaquetário 4 e fator V) e TXA₂ (Oliveira et al., 2010).

Segundo Béréños et al. (2014), as características de tamanho corporal apresentam correlações genéticas positiva de média a alta magnitude, o que faz com que regiões ou cromossomos que explicam grandes quantidades de variação em uma característica expliquem grandes quantidades de variação em outras características. Assim, alguns cromossomos contribuíram significativamente, tanto para peso quanto para comprimento da perna, circunferência da perna, comprimento corporal, altura da cernelha e circunferência torácica, sendo que o cromossomo OAR6 apresentou regiões significativas para as características peso, comprimento da perna e circunferência da perna (Tabela 2 a 4), e o cromossomo 4 com regiões em comum para comprimento corporal, altura da cernelha e para altura da cernelha e circunferência torácica (Tabela 5 a 7).

Diversos trabalhos apresentaram regiões e cromossomos diferentes associados às características de tamanho corporal (Zhang et al., 2013; Riggio et al., 2013; Matika et al., 2016), o que pode ser atribuído ao fato de que a influência genética é explicada por genes diferentes nas diversas raças. Variações na frequência alélica, desequilíbrio de ligação e cobertura de SNPs do chip na raça podem afetar os resultados. Alguns fatores como a metodologia e o número de animais utilizados podem ser responsáveis por estas diferenças.

Conclusões

As regiões identificadas neste estudo poderão auxiliar no processo de seleção para estas características em programas de melhoramento da raça Santa Inês.

Os genes BMP3, SPP1, ABCG2, ALVA, BMPR1B, CCKAR, HGF e CACNA2D1 fazem parte de processos biológicos e funções moleculares importantes para o desenvolvimento corporal, sendo genes candidatos para as características estudadas.

Referências

- AGUILAR, I.; MISZTAL, I.; JOHNSON, D.L.; LEGARRA, A.; TSURUTA, S.; LAWLOR, T.J. Hot topic: A unified approach to utilize phenotypic, full *pedigree*, and genomic information for genetic evaluation of Holstein final score. **J Dairy Sci.**; 93: 743-752. 2010.
- AHMED, M. Y., CHIOZA, B.A., RAJAB, A. et al. Loss of PCLO function underlies ponto cerebellar hypoplasia type III. **Neurology**, 8;84(17) 2015.
- ALEXANDER, L.J.; STEWART, A.F.; MACKINLAY, A.G.; KAPELINSKAYA, T.V.; TKACH, T.M.; GORODETSKY, S.I. Isolation and characterization of the bovine appa-casein gene. **European Journal of Biochemistry**, v.178, p.395-401, 1988.
- BÉRÉÑOS, C.; ELLIS, P.A.; PILKINGTON, J.G.; PEMBERTON, J.M. Estimating quantitative genetic parameters in wild populations: a comparison of *pedigree* and genomic approaches. **Molecular Ecology**, 23, 3434–3451. 2014.
- BROWN, W.M., DZIEGIELEWSKA, K.M., FOREMAN, R.C., SAUNDERS, N.R. Nucleotide and deduced amino acid sequence of sheep serum albumin. **Nucleic Acids Res.** 25;17(24), 1989.
- CASAS, E.; STONE, R. T.; KEELE, J. W.; SHACKELFORD, S. D.; KAPPES, S. M.; KOOHMARAIE, M. A comprehensive search for quantitative trait loci affecting growth and carcass composition of cattle segregating alternative forms of the myostatin gene. **Journal of Animal Science**, v. 79, p. 854-860, 2001.
- DALUISKI, A.; ENGSTRAND, T.; BAHAMONDE, M.E. et al. Bone morphogenetic protein 3 is a negative regulator of bone density. **Nat Genet.** 27:84-8. 2001.
- DENNINGER, K.C.; LITMAN, T.; MARSTRAND, T.; MOLLER, K.; SVENSSON, L.; LABUDA, T.; et al. Kinetics of gene expression and bone remodeling in the clinical phase of collagen-induced arthritis. **Arthritis Res Ther.**; 17: 43.2015.
- DE PEPPEL, J.V.; VAN LEEUWEN, J.; VITAMINA D. And gene networks in human osteoblasts. **Physiol frontal.**5: 137. 2014.

- EAPEN, A., SUNDIVAKKAM, P., SONG, Y., et al. Calcium-mediated Stress Kinase Activation by DMP1 Promotes Osteoblast Differentiation. **The journal of biological chemistry**.285, 47, 36339–36351, 2010.
- FONSECA-ALANIS, M.H.; TAKADA, J.; ALONSO-VALE, M.I.C.; LIMA, F.B. O tecido adiposo como centro regulador do metabolismo. **Arq. Bras. Endocrinol. Metab.**, vol.50,n.2, p 216-229. 2006.
- GUTIÉRREZ-GIL, B.; WILLIAMS, J. L.; HOMER, D.; BURTON, D.; HALEY, C. S.; WIENER, P. Search for quantitative trait loci affecting growth and carcass traits in a cross population of beef and dairy cattle. **Journal of Animal Science**, v. 87, p. 24-36, 2009.
- GUTIERREZ-GIL, B.; JOSÉ ARRANZ, J.; PONG-WONG, R.; GARCIA-GAMEZ, E.; KIJAS, J.; WIENER, P. Application of selection mapping to identify the genomic regions associated with milk production in sheep. **PLoS One**. 9: e94623. 2014.
- KAZMI, H.R.; CHANDRA, A.; NIGAM, J., et al. Polymorphism and Expression Profile of holecystokinin Type A Receptor in Relation to Gallstone Disease Susceptibility. **Biochem Genet**. 54(5):665-75, 2016.
- KIJAS, J.W.; LENSTRA, J.A.; HAYES, B.; BOITARD, S.; PORTO NETO, L. R.; SAN CRISTOBAL, M.; et al. Genome-Wide Analysis of the World's Sheep Breeds Reveals High Levels of Historic Mixture and Strong Recent Selection. **PLoS Biology**, 10(2). 2012.
- KINGSOLVER, J.G.; PFENNIG, D.W. Individual-level selection as a cause of Cope's rule of phyletic size increase. **Evolution**, 58, 1608–1612. 2004.
- KINGSOLVER, J.G.; DIAMOND, S.E. Phenotypic selection in natural populations: what limits directional selection? **The American Naturalist**, 177, 346–357. 2011.
- LIU, R.; SUN, Y.; ZHAO, G.; WANG, F.; WU, D.; ZHENG, M.; et al. Genome-wide association study identifies loci and candidate genes for body composition and quality of meat traits in Beijing-You Chickens. **PLoS One**.;8: e61172. 2013.
- MATIKA, O.; RIGGIO, V.; ANSELME-MOIZAN, M. et al. Association genome-wide reveals QTL for growth, bone and in vivo carcass traits as assessed by computed tomography in Scottish Black sheep. **Genet Sel Evol.**, 48, 11. 2016
- MISZTAL, I.; TSURUTA, S.; STRABEL, T.; AUVRAY, B.; DRUET, T.; LEE, D.H. BLUPF90 and related programs (BGF90). Proc. 7th World Congr. **Genet. Appl. Livest. Prod.**, Montpellier, France. Communication No. 28. 2002.
- MORRISSEY, M.B.; PARKER, D.J.; KORSTEN, P.; et al. The prediction of adaptive evolution: empirical application of the secondary theorem of selection and comparison to the breeder's equation. **Evolution**, 66, 2399–2410. 2012.
- OLESKOVICZ, C.; LISIAUSKAS, S.; ARAGÃO, F.J.L. Expression of human bone morphogenetic protein (BMP-2 and BMP-4) genes in transgenic bovine fibroblasts. **Arq. Bras. Med. Vet. Zootec.**, v.56, p.547-549, 2004.

OLIVEIRA, D.C.; SILVA, R.F.; SILVA, D.J. et al. Resistência à Aspirina: Realidade ou Ficção? **Arq Bras Cardiol**; 95(3): 91-94.2010.

RIGGIO, V.; MATIKA, O.; PONG-WONG, R.; STEAR, M.J.; BISHOP, S.C. Genome-wide association and regional heritability mapping to identify loci underlying variation in nematode resistance and body weight in Scottish Blackface lambs. **Heredity**, 110, 420–429. 2013.

RIMAR, D., ROSNER, I., SLOBODIN G. et al. Semaphorin 3A, a potential immune regulator in familial Mediterranean fever **Clin Exp Rheumatol**. 34(6 Suppl 102):S52-S55, 2016.

RON, N.; KLIGER, D.; FELDMESSER, E. et al. Multiple quantitative trait locus analysis of bovine chromosome 6 in the Israeli Holstein population by a daughter design. **Genetics**. 159:727-735. 2001.

RUBIN, C.J.; MEGENS, H.J.; MARTINEZ BARRIO, A.; MAQBOOL, K.; SAYYAB, S.; SCHWOCHOW, D.; et al. Strong selection signatures in the domestic swine genome. **Proc Natl Acad Sci USA**. 2012;109: 19529- 19536. Doi: 10.1073 / pnas.1217149109

SAHANA, G.; HOGLUND, J.K.; GULDBRANDTSEN, B.; LUND, M.S. The loci associated with adult stature also affect the survival of the calf in cattle. **BMC Genetics** 16: 47, 2015.

SILVA R. M.O., STAFUZZA N.B., FRAGOMENI B.O., et al. (2017) Genome-Wide Association Study for Carcass Traits in an Experimental Nelore Cattle Population. **PLoS ONE** 12(1): e0169860. doi:10.1371/journal.

TETENS, J.; WIDMANN, P.; KÜHN, C.; THALLER, G. A genome-wide association study indicates LCORL / NCAPG as a candidate locus for height of withers in German Warmblood horses. **Anim Genet**. 2013; 44: 467-471. Doi: 10.1111 / idade.12031.

VANRADEN, P.M. Efficient methods to compute genomic predictions. **J Dairy Sci**. 91: 4414±4423.2008.

VISSCHER, P.M.; MACGREGOR, S.; BENYAMIN, B.; et al. Genome partitioning of genetic variation for height from 11,214 sibling pairs. **American Journal of Human Genetics**, 81, 1104–1110. 2007.

WANG, C.; ZÖLLNER, S.; ROSENBERG, N.A. A Quantitative Comparison of the Similarity between Genes and Geography in Worldwide Human Populations. **PLoS Genetics**, 8(8). 2012.

WANG, H.; MISZTAL, I.; AGUILAR, I.; LEGARRA, A, FERNANDO RL, VITEZICA Z, et al. Genome-wide association mapping including phenotypes from relatives without genotypes in a single-step (ssGWAS) for 6-week body weight in broiler chickens. **Front Genet**. 2014; 5:134. 2014.

ZHANG, H.; WANG, Z.; WANG, S.; LI, H. Progress of genome wide association study in domestic animals. **Journal of Animal Science and Biotechnology**, v.3, p.26, 2012.

ZHANG, L.; LIU, J.; ZHAO, F.; REN, H.; XU, L.; LU, J.; et al. Genome-Wide Association Studies for Growth and Meat Production Traits in Sheep. **Plos One**, 8(6). 2013.

Tabela 1. Componentes de variância, estimativas de herdabilidade para características de tamanho corporal em ovinos Santa Inês

| Característica | σ_a^2 | σ_e^2 | h^2 |
|-------------------------|--------------|--------------|-------|
| Peso a idade adulta | 66,46 | 47,36 | 0,58 |
| Comprimento da perna | 4,19 | 7,00 | 0,37 |
| Perímetro da perna | 3,27 | 12,61 | 0,20 |
| Comprimento corporal | 5,02 | 13,29 | 0,27 |
| Altura da cernelha | 2,45 | 7,11 | 0,26 |
| Circunferência torácica | 10,18 | 39,22 | 0,21 |

Tabela 2. Identificação de genes com base no percentual da variância genética aditiva (Var), explicada por janelas de 50 SNPs adjacentes para peso a idade adulto de ovinos da raça Santa Inês

| Cromossomo | Posição (pb) | Gene | Var (%) |
|-------------------|-------------------------|---|--------------------|
| OAR6 | 94573632 | <u>FGF5;LOC105607826; LOC105607825;C6H4orf22;</u> <u>BMP3;LOC105611938;PRKG2;</u> <u>LOC105615531;TRNAW-</u> | 1,08 |
| | 96861918 | <u>CCA;LOC105615532;RASGEF1B;LOC105609536;</u> <u>LOC105609537;LOC105609538;HNRNPD;</u> <u>LOC101108834;LOC101109096;HNRNPDL;ENOPH1</u> | |

Tabela 3. Identificação de genes com base no percentual da variância genética aditiva (Var), explicada por janelas de 50 SNPs adjacentes para comprimento da perna de ovinos da raça Santa Inês

| Cromossomo | Posição (pb) | Genes | Var (%) |
|-------------------|------------------------|---|----------------|
| Oar6 | 124935394- 13523897 | <u>PDE5A;BMPR1B;SNCA;ABCG2;SPP1;TLR10;T</u> <u>LR1;KDR;GNRHR;CSN1S1;CSN2;CSN1S2;CS</u> <u>N3;ALVA;CXCL8;CXCL10;FGF5;HNRNPD;D</u> <u>MP1;SLC2A9;HTT</u> | 1,1 |

Tabela 4. Identificação de genes com base no percentual da variância genética aditiva (Var), explicada por janelas de 50 SNPs adjacentes para perímetro da perna de ovinos da raça Santa Inês

| Cromossomo | Posição (pb) | Genes | Var (%) |
|-------------------|---------------------|--|----------------|
| Oar6 | 45188600-47970141 | <u>SEL1L3;LOC105608876;SMIM20;TRNAA-UGC;TRNAH-AUG;LOC105608875;RBPJ;LOC105615463;LOC105615462;LOC101107788;CCKAR;TBC1D19;TRNAL-UAA;STIM2;LOC105608871;LOC101106847;LOC105615464;LOC101107099</u> | 1,61 |

Tabela 5. Identificação de genes com base no percentual da variância genética aditiva (Var), explicada por janelas de 50 SNPs adjacentes para comprimento corporal de ovinos da raça Santa Inês

| Cromossomo | Posição (pb) | Genes | Var(%) |
|-------------------|-----------------------|---|---------------|
| OAR4 | 35116302- 37850663 | <u>TRNAC-</u> <u>GCA;SEMA3D;LOC105607853;SEMA3A</u> <u>LOC105614979;LOC105614978;SEMA3E;LO</u> <u>C105615152;LOC105614980;PCLO;LOC1011</u> <u>14098</u> | 1,31 |

Tabela 6. Identificação de genes com base no percentual da variância genética aditiva (Var), explicada por janelas de 50 SNPs adjacentes para altura da cernelha de ovinos da raça Santa Inês

| Cromossomo | Posição (pb) | Genes | Var(%) |
|-------------------|-----------------------|--|---------------|
| Oar4 | 38962314 41552192 | <u>HGF;LOC105607851;LOC105614985;LOC105607850;LOC105607849;LOC105614986;LOC105614987;SEMA3C;LOC10102230;LOC10114861;LOC105614988;LOC105614989;LOC101115115;GNAT3;LOC101115786;GNAI1;LOC105613103</u> | 1,91 |
| Oar4 | 35851926- 38337172 | <u>LOC105607853;SEMA3A;LOC105614979;LOC105614978;SEMA3E;LOC105615152;LOC105614980;PCLO;LOC101114098;CACNA2D1;LOC105614984</u> | 1,15 |

Tabela 7. Identificação de genes com base no percentual da variância genética aditiva (Var), explicada por janelas de 50 SNPs adjacentes para circunferência torácica de ovinos da raça Santa Inês

| Cromossomo | Posição (pb) | Genes | Var(%) |
|-------------------|-----------------------|---|---------------|
| Oar13 | 12237644- 15998766 | <u>LOC105610812</u> ; <u>LOC105610813</u> ; <u>LOC105610814</u> ; <u>LOC105610815</u> ; <u>TRNAG-UCC</u> ; <u>TRNAC-ACA</u> ; <u>LOC105616654</u> ; <u>LOC105616653</u> ; <u>LOC105616652</u> ; <u>CELF2</u> ; <u>LOC105616651</u> ; <u>LOC105616650</u> ; <u>LOC105616649</u> ; <u>LOC105616648</u> ; <u>USP6NL</u> ; <u>LOC105611762</u> ; <u>ECHDC3</u> ; <u>LOC105611764</u> ; <u>LOC105611763</u> ; <u>PROSER2</u> ; <u>UPF2</u> ; <u>DHTKD1</u> | 1,13 |
| Oar7 | 14697730- 1946256 | <u>YTHDC2</u> ; <u>PGGT1B</u> ; <u>CCDC112</u> ; <u>LOC105601870</u> ; <u>LVRN</u> ; <u>HRH2</u> ; <u>SFXN1</u> ; <u>ANKDD1B</u> ; <u>CRHBP</u> ; <u>AGGF1</u> ; <u>LOC101108575</u> ; <u>THBS4</u> ; <u>GNPNAT1</u> ; <u>STYX</u> ; <u>GPR137C</u> ; <u>TXNDC16</u> ; <u>PTGER2</u> ; <u>ANKDD1A</u> ; <u>SPG21</u> ; <u>DPP8</u> ; <u>HACD3</u> ; <u>INTS14</u> ; <u>SLC24A1</u> ; <u>DENND4A</u> ; <u>RAB11A</u> ; <u>MEGF11</u> ; <u>MAP2K1</u> ; <u>RPL4</u> ; <u>SMA3</u> ; <u>IQCH</u> | 1,07 |
| Oar4 | 39871116- 42444338 | <u>LOC105614987</u> ; <u>SEMA3C</u> ; <u>LOC101102230</u> ; <u>LOC101114861</u> ; <u>LOC105614988</u> ; <u>LOC105614989</u> ; <u>LOC101115115</u> ; <u>GNAT3</u> ; <u>LOC101115786</u> ; <u>GNAI1</u> ; <u>LOC105613103</u> ; <u>LOC105614015</u> ; <u>LOC105614990</u> ; <u>MAGI2</u> | 1,02 |

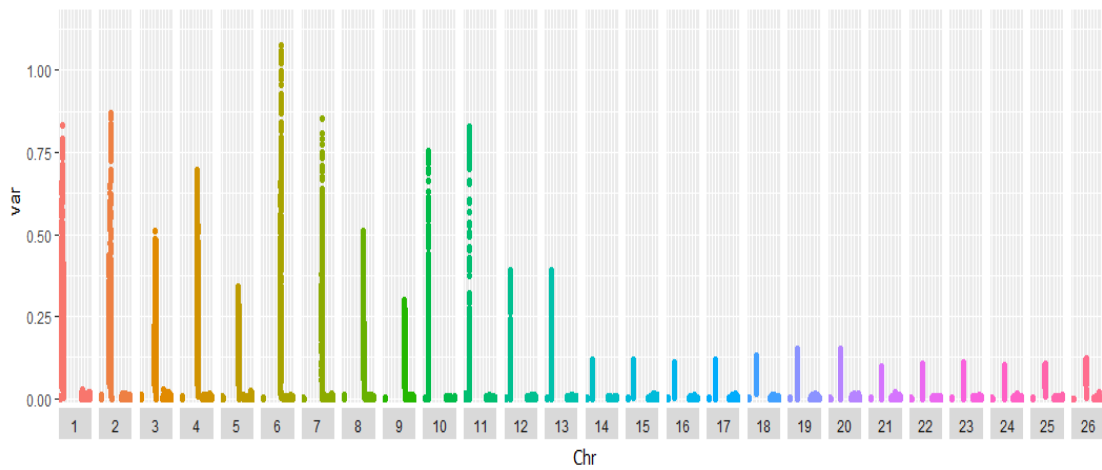


Figura 1. Variância genética aditiva (Var) explicada por janelas de 50 SNPs adjacentes distribuídos por cromossomos (Chr) para peso a idade adulta de ovinos da raça Santa Inês.

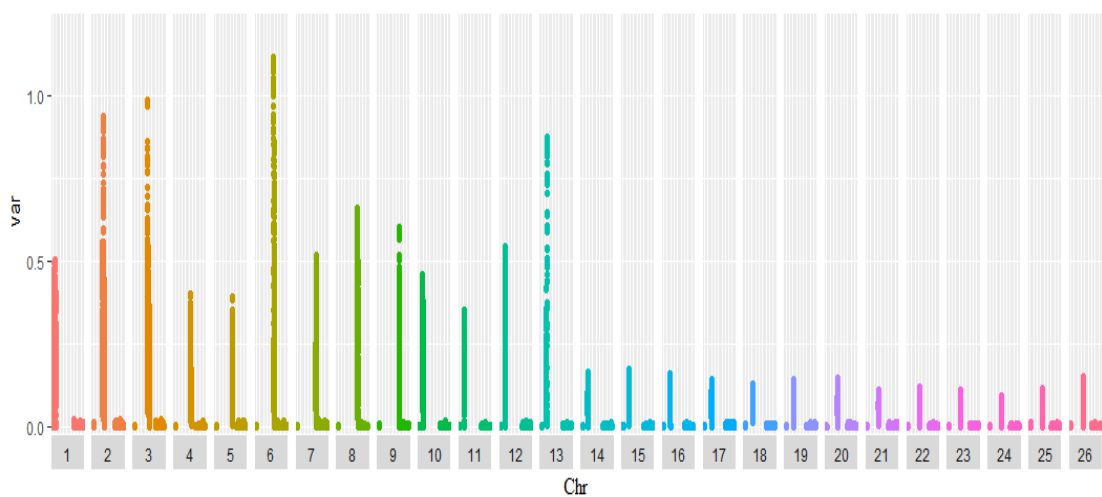


Figura 2. Variância genética aditiva (Var) explicada por janelas de 50 SNPs adjacentes distribuídos por cromossomos (Chr) para comprimento da perna de ovinos da raça Santa Inês.

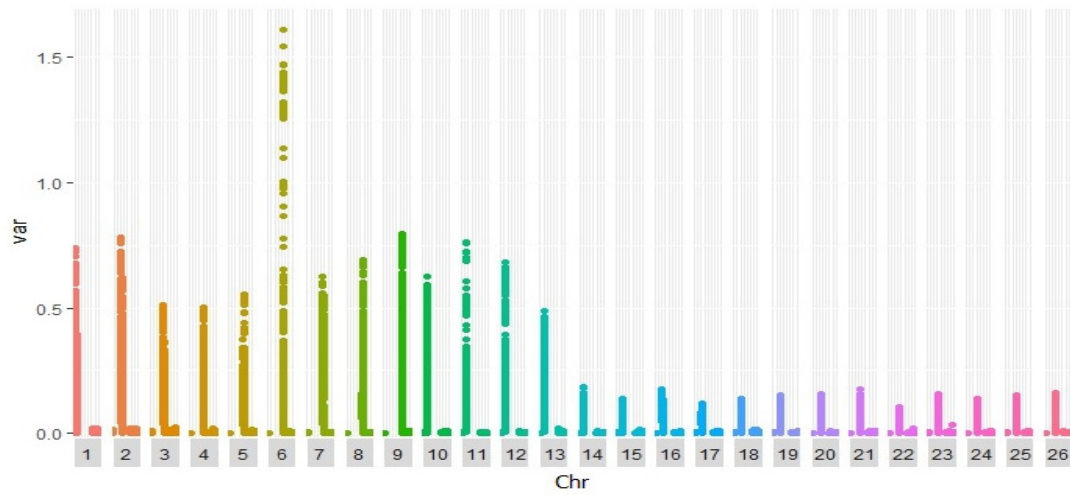


Figura 3. Variância genética aditiva (Var) explicada por janelas de 50 SNPs adjacentes distribuídos por cromossomos (Chr) para perímetro da perna de ovinos da raça Santa Inês.

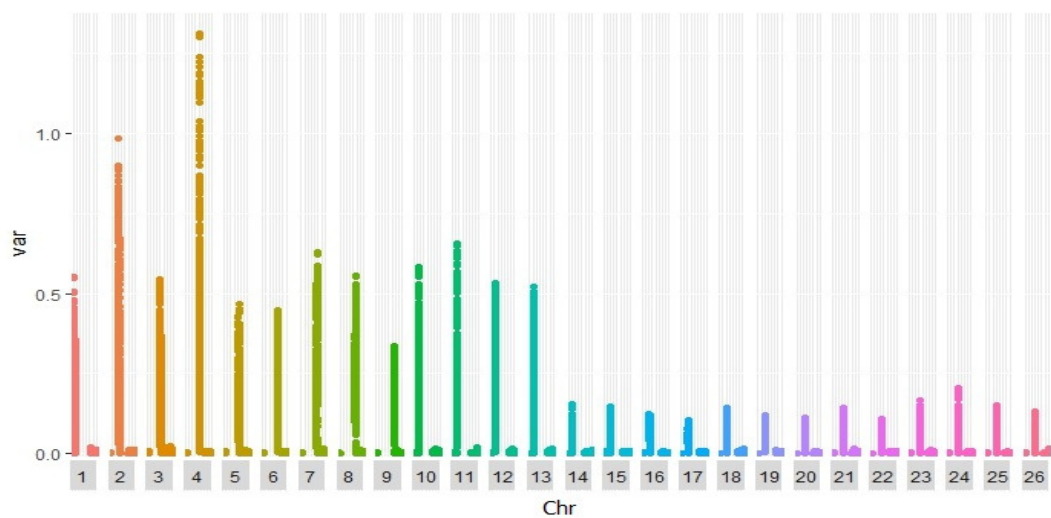


Figura4. Variância genética aditiva (Var) explicada por janelas de 50 SNPs adjacentes distribuídos por cromossomos (Chr) para comprimento corporal de ovinos da raça Santa Inês.

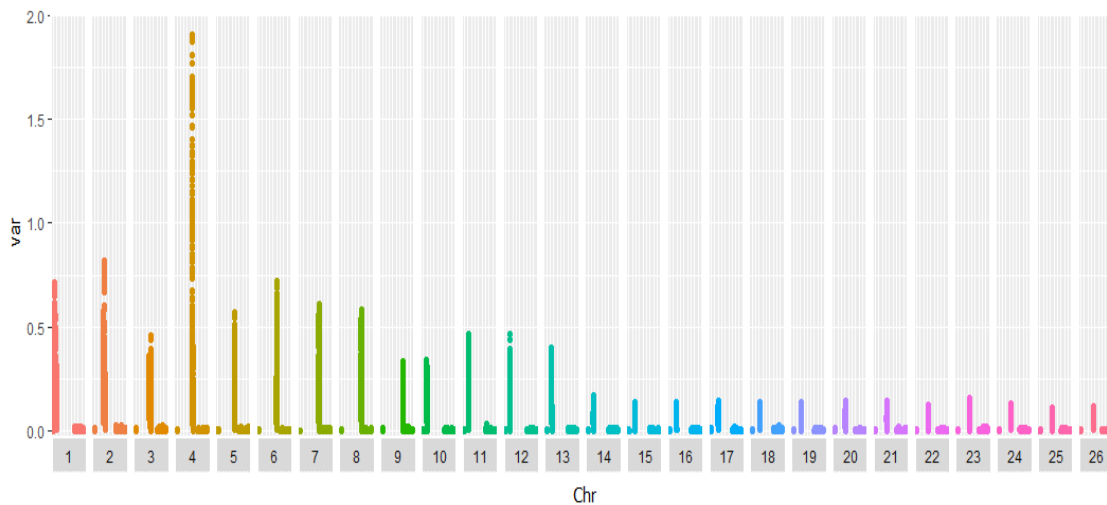


Figura 5. Variância genética aditiva (Var) explicada por janelas de 50 SNPs adjacentes distribuídos por cromossomos (Chr) para altura da cernelha de ovinos da raça Santa Inês.

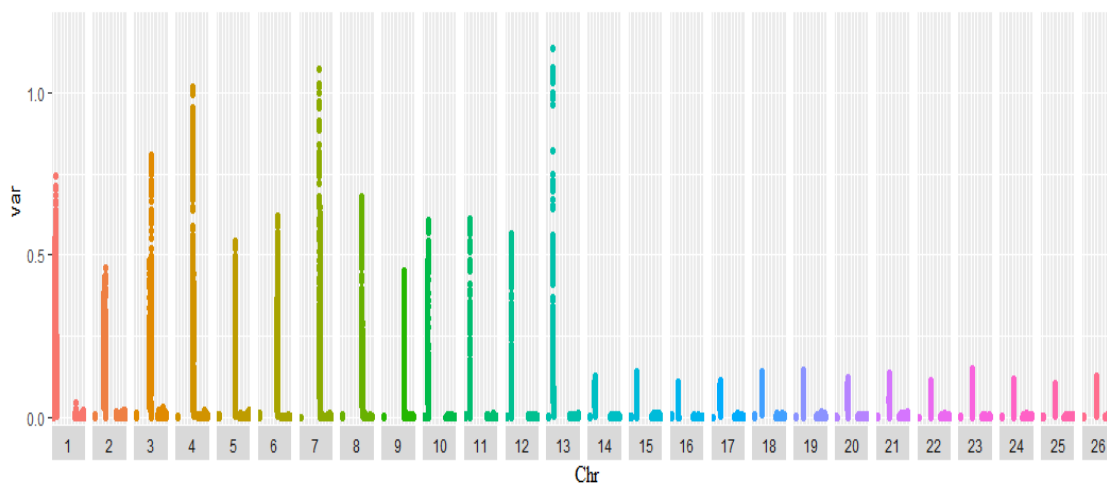


Figura. 6. Variância genética aditiva (Var) explicada por janelas de 50 SNPs adjacentes distribuídos por cromossomos (Chr) para circunferência torácica em ovinos da raça Santa Inês.

5. CONSIDERAÇÕES GERAIS

A diversidade genética da raça Santa Inês deve ser constantemente investigada. Os acasalamentos entre indivíduos menos aparentados devem ser priorizados para garantir a eficiência do programa de melhoramento genético aplicado na população e evitar a depressão endogâmica sobre características de importância econômica.

As características tamanho corporal em ovinos possuem associações genômicas em regiões cromossômicas específicas.

Os resultados da GWAS e a seleção de genes relacionados às características em estudo, servem de base básica para orientar estudos de replicação de acompanhamento, incluindo exploração, análise de rede e verificação funcional nos candidatos, o que pode revelar as mutações causais subjacentes em características de tamanho corporal em ovinos Santa Inês.

6. ANEXOS

Anexo 1

Identificação de genes com base no percentual da variância genética aditiva (Var), explicada por janelas de 50 SNPs adjacentes para peso a idade adulto de ovinos Santa Inês

| Símbolo | Tipo de gene | Descrição |
|------------------------------|--------------------------|--|
| FGF5 | Codificação de proteínas | Factor de crescimento fibroblástico 5 |
| LOC105607826 | Proteína | Lys-63 específicos de deubiquitinase BRCC36-like |
| LOC105607825 | mRNA | Proteína pseudogene FAM60A |
| C6H4orf22 | Proteína | Cromossoma 6, abertos, leitura, Quadro, human, C4orf22 |
| BMP3 | mRNA | Proteína morfogênética óssea 3 |
| LOC105611938 | Rnaseq | Uncharacterized LOC105611938 |
| PRKG2 | Proteína | Proteína quinase, cGMP-dependente, tipo II |
| LOC105615531 | mRNA | Endógeno do grupo K do retrovírus 18 Pol-like |
| TRNAW-CCA | | Transferência de RNA triptofano (anticodon CCA) |
| LOC105615532 | Rnaseq | Uncharacterized LOC105615532 |
| RASGEF1B | mRNA | Membro da família de domínio RasGEF 1B |
| LOC105609536 | Rnaseq | Não classificado LOC105609536 |
| LOC105609537 | Rnaseq | Uncharacterized LOC105609537 |
| LOC105609538 | Rnaseq | Associada à vesícula associada à proteína 714 da membrana |
| HNRNPD | mRNA | Ribonucleoproteína nuclear heterogênea D |
| LOC101108834 | Proteína | Proteína 1 de montagem de nucleosoma semelhante a 1 |
| LOC101109096 | Proteína | Proteína transmembrana induzível por hormônio de crescimento |
| HNRNPDL | mRNA | Ribonucleoproteína nuclear heterogênea D como |
| ENOPH1 | mRNA | Enolase-fosfatase 1 |

Anexo 2

Identificação de genes com base no percentual da variância genética aditiva (Var), explicada por janelas de 50 SNPs adjacentes para comprimento da perna de ovinos Santa Inês

| Símbolo | Tipo de gene | Descrição |
|------------------------|----------------|---|
| | Codificação de | |
| PDE5A | proteínas | Fosfodiesterase 5A |
| BMPR1B | Melhor RefSeq | Receptor de proteína morfogenética óssea tipo 1B |
| SNCA | Melhor RefSeq | Sinucleína alfa |
| ABCG2 | Melhor RefSeq | Subfamília de cassete de ligação de ATP Membro G 2 (Grupo de sangue júnior) |
| SPP1 | Melhor RefSeq | Fosfoproteína 1 segregada |
| TLR10 | Melhor RefSeq | Pedágio como receptor 10 |
| TLR1 | Melhor RefSeq | Pedágio como receptor 1 |
| KDR | Melhor RefSeq | Receptor de domínio de inserto de quinase |
| GNRHR | Melhor RefSeq | Receptor de hormona libertadora de gonadotropina |
| CSN1S1 | Melhor RefSeq | Caseína alfa s1 |
| CSN2 | Melhor RefSeq | Caseína beta |
| CSN1S2 | Melhor RefSeq | Alfa-S2-caseína |
| CSN3 | Melhor RefSeq | Caseína kappa |
| ALVA | Melhor RefSeq | Albumina |
| CXCL8 | Melhor RefSeq | CXC motif chemokine ligand 8 |
| CXCL10 | Melhor RefSeq | Ligante de quimiocina de motivo CXC 10 |
| FGF5 | Melhor RefSeq | Factor de crescimento fibroblástico 5 |
| HNRNPD | MRNA | Ribonucleoproteína nuclear heterogénea D |
| DMP1 | MRNA | Fosfoproteína ácida 1 da matriz da dentina |
| SLC2A9 | Rnaseq | Família de portadores de soluto 2 membros 9 |
| HTT | Melhor RefSeq | Huntingtin |

Anexo3

Identificação de genes com base no percentual da variância genética aditiva (Var), explicada por janelas de 50 SNPs adjacentes para perímetro da perna de ovinos Santa Inês

| Símbolo | Tipo de gene | Descrição |
|--------------------------------------|--------------|--|
| SEL1L3 | mRNA | Membro da família SEL1L 3 |
| LOC105608876 | Rnaseq | Não classificado LOC105608876 |
| SMIM20 | mRNA | Proteína de membrana integral pequena 20 |
| TRNA^A-UGC | | Transferir ARN alanina (anticodon UGC) |
| TRNA^H-AUG | | Transferência ARN histidina (anticodon AUG) |
| LOC105608875 | Rnaseq | Não classificado LOC105608875 |
| RBPJ | mRNA | Proteína de ligação ao sinal de recombinação para a região J kappa da imunoglobulina |
| LOC105615463 | Rnaseq | Não classificado LOC105615463 |
| LOC105615462 | Rnaseq | Uncharacterized LOC105615462 |
| LOC101107788 | Proteína | Pseudogene isopeptidase 2 desumoyling |
| CCKAR | Proteína | Receptor de colecistoquinina A |
| TBC1D19 | MRNA | TBC1 domínio membro da família 19 |
| TRNA^L-UAA | | Transferência de RNA leucina (anticodon UAA) |
| STIM2 | MRNA | Molécula de interação do estroma 2 |
| LOC105608871 | Rnaseq | Uncharacterized LOC105608871 |
| LOC101106847 | Proteína | Antígeno 10 associado ao melanoma |
| LOC105615464 | Rnaseq | Uncharacterized LOC105615464 |
| LOC101107099 | Proteína | 3-oxoacil- [acil-transportador-proteína] sintase, tipo mitocondrial |

Anexo 4

Identificação de genes com base no percentual da variância genética aditiva (Var), explicada por janelas de 50 SNPs adjacentes para comprimento corporal de ovinos Santa Inês

| Símbolo | Tipo de gene | Descrição |
|------------------------------|--------------|---|
| TRNAC-GCA | | Transferência de ARN cisteína (anticodon GCA) |
| SEMA3D | MRNA | Semáfora 3D |
| LOC105607853 | Rnaseq | Não classificado LOC105607853 |
| SEMA3A | MRNA | Semaforina 3 ^a |
| LOC105614979 | Rnaseq | Uncharacterized LOC105614979 |
| LOC105614978 | Rnaseq | Uncharacterized LOC105614978 |
| SEMA3E | MRNA | Semaforina 3E |
| LOC105615152 | proteína | Proteína piccolo-like |
| LOC105614980 | Rnaseq | Uncharacterized LOC105614980 |
| PCLO | MRNA | Piccolo proteína presináptica citomatriz |
| LOC101114098 | proteína | Proteína de transporte vesicular GOT1B pseudogene |

Anexo 5

Identificação de genes com base no percentual da variância genética aditiva (Var), explicada por janelas de 50 SNPs adjacentes para altura da cernelha de ovinos Santa Inês

| Símbolo | Tipo de gene | Descrição |
|------------------------------|--------------|---|
| LOC105607853 | Rnaseq | Não classificado LOC105607853 |
| SEMA3A | mRNA | Semaforina 3 ^a |
| LOC105614979 | Rnaseq | Uncharacterized LOC105614979 |
| LOC105614978 | Rnaseq | Uncharacterized LOC105614978 |
| SEMA3E | mRNA | Semaforina 3E |
| LOC105615152 | proteína | Proteína piccolo-like |
| LOC105614980 | Rnaseq | Uncharacterized LOC105614980 |
| PCLO | mRNA | Piccolo proteína presunáptica citomatríz |
| LOC101114098 | proteína | Proteína de transporte vesicular GOT1B pseudogene |
| CACNA2D1 | mRNA | Subunidade auxiliar do canal de tensão de cálcio alfa2delta 1 |
| LOC105614984 | Rnaseq | Não classificado LOC105614984 |
| HGF | MRNA | Factor de crescimento de hepatócitos |
| LOC105607851 | Rnaseq | Não classificado LOC105607851 |
| LOC105614985 | proteína | Histona-lisina N-metiltransferase tipo SETMAR |
| LOC105607850 | Rnaseq | Membro da família relacionado à caderina 3-like |
| LOC105607849 | Rnaseq | Não caracterizado LOC105607849 |
| LOC105614986 | Rnaseq | Parapente |
| LOC105614987 | Rnaseq | Uncharacterized LOC105614987 |
| SEMA3C | MRNA | Semaforina 3C |
| LOC101102230 | MRNA | Glicoproteína 4 de plaquetas |
| LOC101114861 | MRNA | Glicoproteína 4 de plaquetas |
| LOC105614988 | Rnaseq | Uncharacterized LOC105614988 |
| LOC105614989 | Rnaseq | Uncharacterized LOC105614989 |
| LOC101115115 | MRNA | Glicoproteína plaquetária 4 |
| GNAT3 | proteína | Subunidade da proteína G alfa transducina 3 |
| LOC101115786 | proteína | Cullin-1 pseudogene |
| GNAI1 | MRNA | Subunidade alfa i1 da proteína G |
| LOC105613103 | Rnaseq | Uncharacterized LOC105613103 |

Anexo 6

Identificação de genes com base no percentual da variância genética aditiva (Var), explicada por janelas de 50 SNPs adjacentes para circunferência torácica de ovinos Santa Inês

| Símbolo | Tipo de gene | Descrição |
|------------------------------|---------------|---|
| LOC105614987 | Rnaseq | Uncharacterized LOC105614987 |
| SEMA3C | MRNA | Semaforina 3C |
| LOC101102230 | MRNA | Glicoproteína 4 de plaquetas |
| LOC101114861 | MRNA | Glicoproteína 4 de plaquetas |
| LOC105614988 | Rnaseq | Uncharacterized LOC105614988 |
| LOC105614989 | Rnaseq | Uncharacterized LOC105614989 |
| LOC101115115 | MRNA | Glicoproteína plaquetária 4 |
| GNAT3 | proteína | Subunidade da proteína G alfa transducina 3 |
| LOC101115786 | proteína | Cullin-1 pseudogene |
| GNAI1 | MRNA | Subunidade alfa i1 da proteína G |
| LOC105613103 | Rnaseq | Uncharacterized LOC105613103 |
| LOC105614015 | Rnaseq | Uncharacterized LOC105614015 |
| LOC105614990 | proteína | Fator de transcrição BTF3 pseudogene |
| MAGI2 | MRNA | Guanilato quinase associada a membrana, domínio WW e PDZ contendo 2 |
| YTHDC2 | MRNA | Domínio YTH contendo 2 |
| PGGT1B | MRNA | Proteína geranilgeraniltransferase subunidade tipo I beta |
| CCDC112 | MRNA | Coiled-coil domain contendo 112 |
| LOC105601870 | MRNA | Molécula adaptadora contendo o domínio TIR 2 |
| LVRN | proteína | Laverin |
| HRH2 | MRNA | Histamina receptor H2 |
| SFXN1 | Melhor RefSeq | Sideroflexina 1 |
| ANKDD1B | proteína | Repetição de ankyrina e domínio de morte contendo 1B |
| CRHBP | Melhor RefSeq | Proteína de ligação da hormona libertadora da corticotropina |
| AGGF1 | MRNA | Factor angiogênico com domínios G-patch e FHA 1 |
| LOC101108575 | proteína | S-metilmetionina-homocisteína S-metiltransferase BHMT2 |
| THBS4 | MRNA | Trombospondina 4 |
| GNPNAT1 | MRNA | Glucosamina-fosfato N-acetiltransferase 1 |
| STYX | MRNA | Proteína de interação serina / treonina / tirosina |
| GPR137C | proteína | Receptor acoplado a proteína G 137C |
| TXNDC16 | proteína | Domínio tioredoxina contendo 16 |
| PTGER2 | Melhor RefSeq | Receptor de prostaglandina E 2 |
| ANKDD1A | proteína | Repetição de ankirina e domínio de morte contendo 1ª |
| SPG21 | MRNA | SPG21, maspardin |
| DPP8 | MRNA | Dipeptidil peptidase 8 |

| | | |
|------------------------------|----------|--|
| HACD3 | MRNA | 3-hidroxiacil-CoA desidratase 3 |
| INTS14 | MRNA | Subunidade complexa do integrador 14 |
| SLC24A1 | MRNA | Solute carrier family 24 member 1 |
| DENND4A | MRNA | Domínio DENN contendo 4A |
| RAB11A | MRNA | RAB11A, membro família oncogene RAS |
| MEGF11 | proteína | Domínios múltiplos EGF 11 |
| MAP2K1 | MRNA | Quinase cinase 1 activada por mitogénio |
| RPL4 | MRNA | Proteína ribossomal L4 |
| SMAD3 | mRNA | SMAD membro da família 3 |
| IQCH | proteína | IQ contendo H |
| LOC105610812 | Rnaseq | Não caracterizado LOC105610812 |
| LOC105610813 | Rnaseq | Não classificado LOC105610813 |
| LOC105610814 | Rnaseq | Uncharacterized LOC105610814 |
| LOC105610815 | Rnaseq | Uncharacterized LOC105610815 |
| TRNAG-UCC | | Transferência de RNA glicina (anticodon UCC) |
| TRNAC-ACA | | Transferência ARN cisteína (anticodon ACA) |
| LOC105616654 | Rnaseq | Uncharacterized LOC105616654 |
| LOC105616653 | Rnaseq | Não classificado LOC105616653 |
| LOC105616652 | Rnaseq | Não classificado LOC105616652 |
| CELF2 | mRNA | CUGBP Elav-como membro da família 2 |
| LOC105616651 | mRNA | Não classificado LOC105616651 |
| LOC105616650 | Rnaseq | Não classificado LOC105616650 |
| LOC105616649 | Rnaseq | Não classificado LOC105616649 |
| LOC105616648 | Rnaseq | Não classificado LOC105616648 |
| USP6NL | mRNA | USP6 N-terminal como |
| LOC105611762 | Rnaseq | Uncharacterized LOC105611762 |
| ECHDC3 | MRNA | Enoil-CoA hidratase contendo 3 |
| LOC105611764 | Rnaseq | Não caracterizado LOC105611764 |
| LOC105611763 | Rnaseq | Uncharacterized LOC105611763 |
| PROSER2 | proteína | Prolina e serina rica 2 |
| UPF2 | mRNA | UPF2 regulador de transcritos sem sentido homólogo (levedura) |
| DHTKD1 | mRNA | Desidrogenase E1 e domínio transcetolase contendo 1 |
| LOC105616656 | mRNA | 60S ribossomal protein L30-like |
| SEC61A2 | mRNA | Sec61 translocon subunidade alfa 2 |
| LOC105616655 | mRNA | Não classificado LOC105616655 |
| NUDT5 | mRNA | Nudix hidrolase 5 |
| CDC123 | mRNA | Ciclo de divisão celular 123 |
| LOC105616658 | Rnaseq | Não caracterizado LOC105616658 |
| CAMK1D | mRNA | ID de quinase de proteína dependente de cálcio / calmodulina |