



**UNIVERSIDADE DE BRASÍLIA
FACULDADE DE AGRONOMIA E MEDICINA VETERINÁRIA**

**GENES RELACIONADOS À PROLIFICIDADE E SEU POTENCIAL USO EM
PROGRAMAS DE CONSERVAÇÃO E MELHORAMENTO DE OVINOS.**

THAÍSA SANT'ANNA LACERDA

**TESE DE DOUTORADO
EM CIÊNCIAS ANIMAIS**

**BRASÍLIA/DF
MAIO DE 2016**



**UNIVERSIDADE DE BRASÍLIA
FACULDADE DE AGRONOMIA E MEDICINA VETERINÁRIA**

**GENES RELACIONADOS À PROLIFICIDADE E SEU POTENCIAL USO EM
PROGRAMAS DE CONSERVAÇÃO E MELHORAMENTO DE OVINOS.**

THAÍSA SANT'ANNA LACERDA

**ORIENTADOR: SAMUEL REZENDE PAIVA
CO-ORIENTADOR: ALEXANDRE RODRIGUES CAETANO**

**TESE DE DOUTORADO
EM CIÊNCIAS ANIMAIS**

PUBLICAÇÃO: 163D/2016 BRASÍLIA/DF

**BRASÍLIA/DF
MAIO DE 2016**

REFERÊNCIA BIBLIOGRÁFICA E CATALOGAÇÃO

LACERDA, T. S. Genes relacionados à prolificidade e seu potencial uso em programas de conservação e melhoramento de ovinos. Brasília: Faculdade de Agronomia e Medicina Veterinária, Universidade de Brasília, 2016, 79 p. Tese de Doutorado.

Documento formal, autorizando reprodução desta tese de doutorado para empréstimo ou comercialização, exclusivamente para fins acadêmicos, foi passado pelo autor à Universidade de Brasília e acha-se arquivado na Secretaria do Programa. O autor e o seu orientador reservam para si os outros direitos autorais, de publicação. Nenhuma parte desta dissertação de mestrado pode ser reproduzida sem a autorização por escrito do autor ou do seu orientador. Citações são estimuladas, desde que citada à fonte.

FICHA CATALOGRÁFICA

SL131g SANT'ANNA LACERDA, THAÍSA
GENES RELACIONADOS À PROLIFICIDADE E SEU POTENCIAL
USO EM PROGRAMAS DE CONSERVAÇÃO E MELHORAMENTO DE
OVINOS. / THAÍSA SANT'ANNA LACERDA; orientador
SAMUEL REZENDE PAIVA; co-orientador ALEXANDRE
RODRIGUES CAETANO. -- Brasília, 2016.
79 p.

Tese (Doutorado - Doutorado em Ciência Animal) --
Universidade de Brasília, 2016.

1. Ovis aries, . 2. Paineis de baixa densidade SNP.
3. taxa de ovulação. 4. prolificidade . 5. GDF9,
BMP15, BMP1-b . I. REZENDE PAIVA, SAMUEL , orient.
II. RODRIGUES CAETANO, ALEXANDRE , co-orient. III.
Titulo.

**UNIVERSIDADE DE BRASÍLIA
FACULDADE DE AGRONOMIA E MEDICINA VETERINÁRIA**

**GENES RELACIONADOS À PROLIFICIDADE E SEU POTENCIAL USO EM
PROGRAMAS DE CONSERVAÇÃO E MELHORAMENTO DE OVINOS.**

THAÍSA SANT'ANNA LACERDA

TESE DE DOUTORADO SUBMETIDA AO PROGRAMA DE
PÓS-GRADUAÇÃO EM CIÊNCIAS ANIMAIS, COMO
PARTE DOS REQUISITOS NECESSÁRIOS À OBTENÇÃO
DO GRAU DE DOUTOR EM CIÊNCIAS ANIMAIS.

APROVADA POR:

SAMUEL REZENDE PAIVA, Pós-doutorado (EMBRAPA)
(ORIENTADOR)

CAROLINA MADEIRA LUCCI, Doutorado (UNIVERSIDADE DE BRASILIA)
(EXAMINADOR INTERNO)

ALEXANDRE FLORIANI RAMOS, Doutorado (EMBRAPA RECURSOS GENÉTICOS E
BIOTECNOLOGIA - CENARGEN)
(EXAMINADOR EXTERNO)

PATRÍCIA IANELLA, Doutorado (EMBRAPA RECURSOS GENÉTICOS E
BIOTECNOLOGIA - CENARGEN)
(EXAMINADOR EXTERNO)

OLIVARDO FACÒ, Doutorado (EMBRAPA CAPRINOS E OVINOS)
(EXAMINADOR EXTERNO- SUPLENTE)

BRASÍLIA/DF, 4 DE MAIO DE 2016

AGRADECIMENTOS

A DEUS, que foi e sempre será minha força e minha motivação, por ter me capacitado e por me dar força e Fé nos momentos de grandes medos e incertezas, todas honras e glórias.

Ao meu orientador Dr. Samuel Rezende Paiva pela orientação, incentivo, compreensão, paciência, amizade, reconhecimento pela competência e pelos ensinamentos que me proporcionou.

Ao meu co-orientador Dr. Alexandre Rodrigues Caetano pelo suporte, apoio e ajuda toda vez que solicitado.

A professora Dra. Concepta Margaret McManus Pimentel pelos ensinamentos, disponibilidade e auxílio neste trabalho.

Ao pesquisador Dr Michel Eduardo Beleza Yamagishi pelo grande auxílio e ensinamentos na parte de Bioinformática.

Ao pesquisador Dr Harvey Blackburn pela oportunidade de trabalhar e disponibilizou amostras de ovinos do banco de germoplasma do NCGRP.

A minhas amigas: Danielle Paiva, Bruna Ohse, Natalia Lamas, Dione Mendes, Lilian Cavalcanti, Ana Ledo, Eleonora Barbosa e Mirella Ortiz pela ajuda, conselhos, momentos de descontração e pela maravilhosa amizade que me proporcionaram durante a vida acadêmica.

À Embrapa Recursos Genéticos e Biotecnologia por toda infraestrutura, apoio, suporte financeiro e oportunidade no desenvolvimento deste trabalho.

À Universidade de Brasília e ao Programa de Pós-Graduação em Ciências Animais pelo suporte e oportunidade de realização do curso.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão da bolsa de formação durante o curso de doutorado.

Ao Programa Ciência sem Fronteiras-CNPq pela bolsa de Doutorado-Sanduíche, que foi realizado no *National Center for Genetic Resources Preservation* (NCGRP/ARS), em Fort Collins, Colorado nos Estados Unidos da América.

A todos colegas do Laboratório de Genética Animal.

A todos meus amigos e família

E a todos aqueles que, de alguma forma, contribuíram para a realização deste trabalho o meu MUITO OBRIGADO.

ÍNDICE

RESUMO	viii
ABSTRACT.....	ix
LISTA DE FIGURAS	x
ÍNDICE DE TABELAS	xi
LISTA DE SÍMBOLOS E ABREVIACÕES.....	xii
CAPÍTULO 1	1
1 INTRODUÇÃO.....	2
1.1 Problemática e relevância.....	2
1.2 Objetivos.....	5
2 REVISÃO DE LITERATURA	6
2.1 Breve estado da arte da ovinocultura no Brasil.	6
2.2 Breve estado da arte do programa de conservação de recursos genéticos animais no Brasil.	7
2.3 Breve estado da arte do programa de melhoramento de ovinos no Brasil.	9
2.4 Prolificidade e taxa de ovulação em ovinos.	10
2.5 Genes associados a taxa de ovulação.	11
2.6 Impacto das novas tecnologias genômicas em características controladas por poucos genes.	14
2.6.1 Sequenciamento de nova geração.....	15
2.6.2 Painéis de genotipagem de SNPs de baixa densidade	16
2.6.3 Estudos de associações de genômica ampla (GWAS).....	17
3 REFERÊNCIAS BIBLIOGRÁFICAS	19
CAPÍTULO 2 - SELEÇÃO ASSISTIDA POR MARCADOR DE BASE ÚNICA (SNP) EM OVINOS DOS NÚCLEOS DE MELHORAMENTO GENÉTICO PARTICIPATIVO DA RAÇA MORADA NOVA. *	25
RESUMO	26
ABSTRACT	27
1 INTRODUCTION.....	28
2 MATERIAL AND METHODS.....	30
3 RESULTS.....	32

4 DISCUSSION.....	35
5 CONCLUSION	37
6 REFERENCES	38
CAPÍTULO 3 - MINERAÇÃO DE POLIMORFISMOS DE BASE ÚNICA (SNPs) EM GENES RELACIONADOS À TAXA DE OVULAÇÃO EM OVINOS	40
RESUMO	41
ABSTRACT	42
1 INTRODUÇÃO.....	43
2 MATERIAIS E MÉTODOS.....	45
2.1 Material biológico e amostragem	45
2.2 Montagem das sequências completa dos genes candidatos selecionados utilizando a População 1	48
2.3 Identificação de SNPs na População 1 para os três genes selecionados.	49
2.4 Amplificação, sequenciamento e alinhamento das sequências gênicas do exon 2 do gene <i>GDF9</i> . (População 2).....	50
2.5 Análises estatísticas	50
3 RESULTADOS	51
3.1 Montagem e alinhamento da sequência completa dos genes <i>BMP15</i> , <i>BMPRI-b</i> e <i>GDF9</i>	51
3.2 Identificação de SNPs nos genes: <i>GDF9</i> , <i>BMPRI-b</i> e <i>BMP15</i> para População 1	52
3.3 Estimativas de haplótipos formados pelos SNPs identificados nos genes <i>GDF9</i> , <i>BMP15</i> e <i>BMPRI-b</i>	60
3.4 Estimativas de haplótipos no exon 2 do gene <i>GDF9</i> para População 2.	64
4 DISCUSSÃO	66
5 CONCLUSÃO.....	69
6 REFERÊNCIAS BIBLIOGRÁFICAS	70
CAPITULO 4	73
CONSIDERAÇÕES FINAIS E CONCLUSÕES	74
REFERÊNCIAS BIBLIOGRÁFICAS.....	76
ANEXO	77

RESUMO

GENES RELACIONADOS À PROLIFICIDADE EM USO DE PROGRAMAS DE CONSERVAÇÃO E MELHORAMENTO DE OVINOS. Thaísa Sant'Anna Lacerda¹
Samuel Rezende Paiva, Ph.D.²

1 – Doutoranda na UnB, Brasília/DF

2 – Pesquisador Embrapa Labex EUA – Secretaria Relações Internacionais

A prolificidade (número de cordeiros nascidos por ovelhas expostas à reprodução) está relacionada à taxa de ovulação e pode ser considerada como um parâmetro importante no aumento da produtividade. Atualmente, a fenotipagem dessa característica só pode ser avaliada em fêmeas em idade reprodutiva e, em geral, apresenta um custo relativamente alto para espécie. No entanto, esta é uma característica controlada por poucos genes de forma que pode ser factível a implementação de esquemas de seleção por marcadores moleculares. Três genes de efeito principal tem sido considerados para explicar grande parte dos fenótipos de partos múltiplos: *GDF9*, *BMP15* e *BMPRI-b*. O presente trabalho foi executado de forma a aumentar o conhecimento dos polimorfismos existentes nestes genes em raças localmente adaptadas de ovinos bem como avaliar a possibilidade de aplicar conhecimento adquirido em programas de conservação e melhoramento no Brasil. No primeiro experimento, 252 animais da raça Morada Nova provenientes de testes de desempenho do Programa de Melhoramento genético participativo da raça foram genotipados para presença ou ausência do alelo *FecG^E* do gene *GDF9*, inicialmente descrito apenas na raça Santa Inês. A referida mutação foi identificada em alta frequência na raça Morada Nova e ela não aparenta estar associada a características de produção/crescimento usadas nos testes de desempenho da raça Morada Nova. No segundo experimento, sequências completas dos genes: *BMP15*, *BMPRI-b* e *GDF9* foram obtidas por dados de sequenciamento de nova geração de 75 ovinos de diferentes raças provenientes do repositório do Consórcio Internacional do Genoma Ovino (ISGC). Foram identificados um total de 1610 SNPs a partir de dados de sequenciamento de nova geração de 75 genomas ovinos. Pelo menos três SNPs foram identificados na região de exon dos genes *BMPRI-b* e *BMP15*, que resultaram em alterações não conservativa de aminoácidos e, portanto, são candidatos imediatos a serem testados em raças com históricos de partos múltiplos. A identificação e seleção destes SNPs serão usadas para compor um painel de baixa densidade para a prolificidade e, depois de validados em raças prolíficas com histórico de parto simples e múltiplos, poderão ser usados para otimizar programas de conservação e melhoramento de ovinos.

Palavras Chaves: *GDF9*, *BMP15*, *BMPRI-b*, *Ovis aries*, prolificidade, taxa de ovulação, SNP, recursos genéticos animais.

ABSTRACT

GENES ASSOCIATED WITH PROLIFICACY FOR USE IN SHEEP CONSERVATION AND BREEDING PROGRAMMES. Thaísa Sant'Anna Lacerda¹
Samuel Rezende Paiva, Ph.D².

1 – School of Agronomy and Veterinary Medicine-UnB, Brasília/DF

2 –Embrapa Labex EUA

The prolificacy (number of lambs per sheep exposed to reproduction) is mainly related to ovulation rate and can be considered as an important parameter in sheep productivity. This phenotype can only be evaluated in reproductive age female and still has a reasonable cost. On the other hand, it is known this trait is controlled by few genes and it might be feasible to implement marker assisted selection scheme. Three major genes have been recognized to explain phenotypes related to litter size: *GDF9*, *BMP15* and *BMPRI-b*. The objective of the present study was to increase the knowledge of existing polymorphisms in genes related to prolificacy in locally adapted sheep breeds and evaluate its use in conservation and breeding programs in Brazil. In the first experiment it was genotyped, for the *FecG^E* mutation (*GDF9*), 252 Morada Nova animals who participated in breed performance tests. The *FecG^E* mutation is not restricted to Santa Ines breed and it was found in high frequency in Morada Nova sheep. In addition, this mutation does not appear to be associated with production traits used on performance tests of Morada Nova sheep breed. The second experiment was analyze the complete sequence of genes: *BMP15*, *BMPRI-b* and *GDF9* from 75 sheep of different breeds from the International Consortium repository Genome Ovine (ISGC). It was identified a total of 1610 SNPs from 75 sheep genomes. Three SNPs from *BMP15*, *BMPRI-b* genes change in non-conservative amino acid. The identification and selection of these SNPs will be used to develop a low density panel for prolificacy that might be applied to conservation and breeding programs.

Keywords: *GDF9*, *BMP15*, *BMPRI-b*, *Ovis aries*, prolificacy, ovulation rate, SNP, animal genetic resources

LISTA DE FIGURAS

CAPÍTULO 1

1- Diferentes alelos nos genes *GDF9*, *BMP15* e *BMPRI-b* envolvidos na regulação do desenvolvimento folicular ovariano resultando em fenótipos de hiperprolificidade e infertilidade em diferentes raças ovinas. Fonte: (adaptada- Juengel et al., 2013).----- 13

CAPÍTULO 2

1- Logistic regression of genotypes *FecG^E/FecG^E* (1), *FecG^E/FecG⁺* (2) and *FecG⁺/FecG⁺* (3) for Morada Nova animals classified as Regular/inferior in performance test (grp=0)----- 34

ÍNDICE DE TABELAS

CAPÍTULO 1

- 1- Mutações encontradas em genes associados à taxa de ovulação relacionadas com fenótipos de prolificidade em ovinos----- 12

CAPÍTULO 2

- 1- *FecG^E* allele and genotype frequencies by farm for Morada Nova sheep from Ceará (N= 140) and São Paulo (N= 112) States, Brazil.----- 33
- 2- *FecG^E* genotypic and allelic frequencies in Morada Nova rams in Ceará (N= 122) and São Paulo (N= 17) States according to performance test classification.----- 34

CAPÍTULO 3

- 1-Os 75 Animais de 39 raças e duas espécies selvagens que serão usados para a análise de sequências dos genes associados à alta taxa de ovulação, provenientes do repositório do Consórcio Internacional do Genoma Ovino (ISGC) ----- 47
- 2- Resultados do alinhamento das sequências REFIGSC contra os cromossomos da sequência referência (REFOar3.1) depositada no Genbank (refSeq).----- 53
- 3- Distribuição por classe e localização dos SNPs encontrados após busca de variantes nos genes *GDF9*, *BMPRI-B* e *BMP15* na População1 (n=75). ----- 55
- 4- Sítios polimórficos (36) encontrados nas regiões de exon dos genes: *GDF9*, *BMPRI-b* e *BMP15* na Pop1 (n=75).----- 56
- 5- Sítios polimórficos encontrados em região de exon no gene *GDF9* e sua frequência absoluta por raça. ----- 57
- 6- Sítios polimórficos encontrados no exon do gene *BMPRI-b* e sua frequência absoluta por raça. ----- 58
- 7- Sítios polimórficos encontrados no exon do gene *BMP15* e sua frequência absoluta por raça. ----- 59
- 8- Segunda estimativa de haplótipos utilizando SNPs encontrados nos exons dos genes (com mutação não conservativa) e suas respectivas mutações e frequências.----- 62
- 9- Caracterização dos SNPs encontrados via sequenciamento do exon 2 do gene *GDF9* na Pop2. ----- 64
- 10- 30 Haplótipos estimados na Pop2 (N=439) em uma região do exon 2 no gene *GDF9* ----- 65

ANEXO

- 1A- Haplótipos estimados para os genes *GDF9*, *BMP15*, *BMPRI-b* e sua distribuição por raças na Pop1.----- 78

LISTA DE SÍMBOLOS E ABREVIACÕES

%Genot- Percentual de indivíduos genotipados

A - Adenina

AA- Aminoácido

ARCO - Associação brasileira de criadores de ovinos

BCF- *Binary Call Format*

BLAST - *Basic Local Alignment Search Tool*

BMPs- *Bone morphogenetic proteins*

BMP15- Bone morphogenetic proteins factor 15

BMPRI-b- Bone morphogenetic proteins receptor factor 1b.

BWA- *Burrows-Wheeler Aligner*

C - Citosina

CENARGEN- Embrapa Recursos Genéticos e Biotecnologia.

CNPCO- Embrapa Caprinos e Ovinos

CPATC- Embrapa Tabuleiros Costeiros

Cr.- Cromossomo

DL-Desequilíbrio de Ligação

DH- Diversidade Haplotípica

EHW- Equilíbrio de Hardy-Weinberg

EMBRAPA - Empresa Brasileira de Pesquisa Agropecuária

EMEPA-PB - Empresa estadual de pesquisa agropecuária da Paraíba

FAO- *Food and Agriculture Organization of the United Nations*

FecB- Fecundity gene BMPRI-b

FecG- Fecundity gene GDF9

FecX- Fecundity gene Chromosome X

FSH- Hormônio Folículo Estimulante

G - Guanina

GDF9- Growth and differentiation factor

GEBV- *Genomic Estimated Breeding Value*

GENECOC - Programa de melhoramento genético de caprinos e ovinos de corte

GWAS- *Genome Wide Association Study*

Het. - Heterozigosidade

HetEsp- Heterozigozidade Esperada
HetObs. – Heterozigozidade Observada
IBGE - Instituto Brasileiro de Geografia e Estatística
IGV- Integrative Genome Viewer
InDel- Inserção ou deleção de nucleotídeos
ISGC- *International Sheep Genome Consortium*
LGA- Laboratório de Genética Animal
LH- Hormônio Luteinizante
MAF-*Minor allele Frequency*- Frequência do alelo menor
MALDI-TOF - *Matrix-assisted laser desorption/ionization*
MAS- *Marker Assisted Selection* - Seleção Assistida por Marcadores
NCBI - *National Center for Biotechnology Information*
NCGRP- *National Center Genetic Resources Preservation*
NGS- *Next Generation Sequence*
OCL-Ovinos Crioula Lanada
OMN-Ovinos Morada Nova
OPPV- vírus da pneumonia progressiva ovina
OSI- Ovinos Santa Inês
OVINOPLUS -Programa de Avaliação Genética de Ovinos
P-value- Valor $-p$, valor de significância estatística
pb - Pares de bases
PROMOVI - Programa de melhoramento genético de ovinos
QTL - *Quantitative trait loci*
REFISGC-Referência gênica baseada nos 75 animais do repositório do ISGC
REFOar3.1- Referência gênica baseada na montagem do genoma de ovino versão 3.1.
SANTAGEN - Programa de melhoramento genético de ovinos Santa Inês
SNP - Single Nucleotide Polimorphism
T - Timina
TGF- β - Transforming growth factor beta
VCF- Variant Call Format file
UFPI-Universidade Federal do Piauí
USP-Universidade de São Paulo

CAPÍTULO 1

1 INTRODUÇÃO

1.1 Problemática e relevância

A domesticação dos ovinos resultou em uma seleção de raças específicas para a produção de lã, carne e pele em diferentes ambientes. O resultado dessa pressão de seleção e da deriva genética acarretou a formação de diversas raças com uma grande amplitude de fenótipos e, segundo Primo (2004), uma alta capacidade de adaptação a todos os continentes e ecossistemas existentes. Estudos genéticos para caracterizar as consequências do processo de domesticação foram realizados em uma amostragem de 74 raças de ovinos ao redor do mundo e mostraram que os ovinos contêm maior diversidade genética e maior tamanho efetivo populacional ancestral quando comparados com outras espécies domesticadas (Kijas et al., 2012). O processo de domesticação da espécie ocorreu a partir de uma base genética mais ampla (Kijas et al., 2012) que realça o potencial de seleção para características de interesse econômico.

O rebanho mundial de ovinos é hoje representado por volta de 1 bilhão de cabeças, destacando a China, Austrália, Índia, Iran e Sudão como os maiores produtores (FAO, 2013). O Brasil hoje ocupa o 18º lugar representado por uma população de 17,291 milhões de animais (IBGE, 2013). Os ovinos foram introduzidos no Brasil há aproximadamente 500 anos vindos, principalmente, do Paraguai e Argentina no período do descobrimento os quais se adaptaram às condições locais de cada região (McManus et al., 2010). Na última década, a ovinocultura aumentou sua participação no agronegócio total brasileiro, e ainda se estima que esse crescimento se mantenha em expansão (IBGE, 2013). No último século, houve um aumento de 24,4% no tamanho populacional de ovinos no Brasil, porém esse crescimento ainda foi aquém quando comparado com outras espécies de animais de produção. Adicionalmente, a produtividade da ovinocultura ainda é considerada baixa (Hermuche et al., 2013a).

Historicamente, a produção de ovinos no Brasil foi vista como uma cultura secundária, principalmente representada por produção de subsistência de pequenos agricultores (Hermuche et al., 2013a; Lobo, 2006; McManus et al., 2010). De maneira geral, os principais gargalos para a ovinocultura no Brasil podem estar associados aos seguintes fatores: problemas sanitários, manejo reprodutivo deficiente, inexistência de um programa amplo de melhoramento genético dos rebanhos e baixos investimentos em pesquisa e tecnologias de produção.

A necessidade por ações dos programas de melhoramento e conservação de ovino no Brasil surgiu com a alta demanda por produtos. Pesquisas voltadas à essas áreas visam, em médio e longo prazo, a avaliação e identificação de animais superiores que, em último caso, poderão ser importantes para aumentar a produtividade dos rebanhos (Souza et al., 2009; McManus et al., 2010; Silva et al., 2011; Kijas et al., 2012). Um dos gargalos para expansão dos programas de melhoramento deve-se ao pequeno interesse e a participação dos produtores. O que pode ser devido ao desconhecimento da importância do melhoramento genético para o aumento da produtividade (Lobo, 2006) por parte dos mesmos. Dentre os programas de melhoramento de ovinos existentes atualmente no Brasil pode se destacar o programa de melhoramento genético GENECOC como o núcleo participativo de ovinos da raça Morada Nova coordenado pela Embrapa Caprinos e Ovinos desde 2008.

Após a crise internacional da lã na década de 90, cresceu o interesse em raças deslanadas pelos criadores e associações, de forma que a região nordeste tornou-se a maior produtora de ovinos do país (IBGE, 2013). No entanto, as raças deslanadas apresentam baixos índices de produtividade (Sousa & Leite, 2000), taxas de crescimento mais lento e qualidade da carne inferior quando comparadas as raças lanadas (McManus, et. al., 2010). Além do mais, cruzamentos direcionados entre as principais raças localmente adaptadas do Nordeste com raças especializadas vêm ocorrendo com grande frequência reduzindo a adaptação desses animais aos sistemas de produção locais (Paiva et al., 2005). Tais fatores podem ser controlados por investimentos em pesquisas voltados para o melhoramento na produção e o uso das raças apropriadas em sistema de produção adequado. A implementação dos programas de melhoramento e conservação são estratégias de longo prazo imprescindíveis para minimizar os riscos expostos acima bem como qualificar os grupos genéticos para atenderem as demandas de produção do país. Uma estratégia alternativa de curto prazo poderia ser a exploração de características de interesse produtivo controladas por poucos genes. Essas características podem ser selecionadas de maneira rápida e acurada, via marcadores moleculares, com ganhos reais para os sistemas de produção.

A prolificidade, número de cordeiros nascidos por ovelhas expostas à reprodução, se corretamente manejada, pode ser considerada como um parâmetro importante no aumento da escala de produção (Lôbo et al., 2011). Como a prolificidade é dependente da taxa de ovulação (Galloway et al., 2000) esse seria um dos entraves para seleção massal dessa característica, já que a fenotipagem pode ser avaliada apenas em fêmeas com idade reprodutiva e possui um custo relativamente elevado quando comparado ao valor do animal comercializado no país. Porém estudos foram realizados em genes relacionados à taxa de ovulação de ovinos e alguns candidatos para seleção assistida por Marcadores (*MAS- Marker Assisted Selection*) foram localizados nos cromossomos X (*BMP15*), 5 (*GDF9*) e 6 (*BMPRI-b*) (Otsuka et al., 2011).

Nos últimos cinco anos as tecnologias de sequenciamento de nova geração e de genotipagem em larga escala avançaram de forma exponencial gerando grandes impactos em pesquisas com a espécie ovina (Heaton et al., 2013; Kijas et al., 2012). Com isso, estudos de associação genômica ampla (*Genome Wide Association Study - GWAS*) para características controladas por poucos genes, passaram a ser realizados de maneira satisfatória, em um curto período de tempo quando comparado as tecnologias anteriores como microssatélites ou re-sequenciamento utilizando metodologia de Sanger.

A partir dessa realidade tecnológica, bem como dos cenários para produção de ovinos no Brasil (Hermuche et al., 2013a) o presente trabalho foi proposto para aumentar o conhecimento dos polimorfismos existentes em genes relacionados a prolificidade em raças localmente adaptadas de ovinos bem como aplicar parte dessas tecnologias em programas de conservação e melhoramento no Brasil.

1.2 Objetivos

1.2.1 Objetivo Geral

Gerar informações genômicas e adaptar tecnologias para auxiliar programas de melhoramento e conservação de ovinos em genes relacionados à taxa de ovulação.

1.2.2 Objetivos específicos

- Fazer levantamento da frequência da mutação *FecG^E* (gene *GDF9*) em animais do programa de melhoramento participativo da raça Morada Nova
- Analisar o potencial de se incorporar a mutação *FecG^E* (gene *GDF9*) em testes de desempenho da raça Morada Nova.
- Identificar marcadores moleculares SNP nos principais genes (*BMP15*, *BMPRI-b* e *GDF9*) relacionados à taxa de ovulação na espécie ovina por meio de dados de sequenciamento de nova geração.

2 REVISÃO DE LITERATURA

2.1 Breve estado da arte da ovinocultura no Brasil.

A ovinocultura constitui uma importante atividade econômica regional para o país, apresentando um crescimento da produção nestas últimas décadas. Segundo dados publicados pelo IBGE (2013), o rebanho brasileiro é representado por mais de 17 milhões de cabeças distribuídas por todo o país de acordo com as adaptações das raças às regiões, destacando as regiões Nordeste e Sul com maior número de cabeças no Brasil. A criação de ovinos na região Sul é baseada em raças de ovinos lanados adaptados ao clima subtropical com aptidão mista para a produção de carne e lã. No Nordeste do Brasil a produção é voltada para raças de ovinos especializadas para corte representado por raças deslanadas adaptadas ao clima tropical da região e com alta rusticidade e capacidade adaptativa às condições peculiares da região; além de servir como fontes de proteína para as populações locais (McManus, et al., 2010; McManus et al., 2014).

A produção de carne é o maior insumo da ovinocultura na região Nordeste e tem sido observado um contínuo crescimento no número efetivo de ovinos na região. Com a crise internacional da lã na década de 1990, a demanda por raças especializadas na produção de carne em todo Brasil estimulou a criação de ovinos deslanados principalmente na região Nordeste (Viana, 2008). Assim, a exploração econômica pelo uso de raças especializadas e adaptadas ao ambiente da região, e sistemas de produção, estimulou o melhoramento genético e técnicas de manejo, aumentando a produtividade dos rebanhos no Nordeste brasileiro. O Nordeste brasileiro é considerado atualmente o novo centro produtor de ovinos (Viana, 2008). Ao longo dos últimos anos, tem sido observado exigências de carcaças de melhor qualidade por

consumidores brasileiros e melhoria na regulamentação do Governo Federal para a comercialização de carnes. E como isso, observa-se também o aumento no número de pesquisas voltadas para o melhoramento genético e manejo de rebanhos especializados, maior interesse de produtores e conseqüentemente melhor qualidade dos animais produzidos (Hermuche et al., 2013a; Simplício & Simplício, 2006).

2.2 Breve estado da arte do programa de conservação de recursos genéticos animais no Brasil.

A introdução de animais de produção no Brasil teve início com colonizadores portugueses que os trouxeram em seus navios (McManus et al., 2010). As raças dos animais chegadas nesta época eram originadas da Europa e África (Primo, 2004), e com o passar dos anos passaram por processos de seleção natural e adaptação resultando em diferentes raças localmente adaptadas na maioria dos diversos ecossistemas brasileiros (Hermuche et al., 2013b; Paiva et al., 2005). Não existem registros específicos de ovelhas introduzidas diretamente para o Brasil vindos da Europa (Rodero et al., 1992), mas sim ovinos introduzidos no Brasil, principalmente, vindos do Paraguai e Argentina no período do descobrimento os quais se adaptaram às condições locais de cada região (McManus et al., 2010).

Animais de raças localmente adaptadas desenvolveram características únicas como a rusticidade, prolificidade e, resistência às doenças e parasitas, com uma vasta diversidade genética, que são considerados hoje, um importante material genético para a conservação (Mariante et al., 2011). A conservação dessas raças permitirá, no futuro, sua utilização em cruzamentos para melhorar a resistência e adaptação de outras raças a condições ambientais desfavoráveis contribuindo para a melhoria na produção de animais doméstico no país (Mariante et al., 2011).

Os dois últimos séculos foram marcados por um aumento na importação de raças exóticas e produção de raças comerciais substituindo as raças localmente adaptadas, e impactando diretamente na diminuição da variabilidade genética culminando em uma situação de ameaça de extinção destas raças. Desta forma, programas de conservação são cada vez mais necessários para ampliar a capacidade de gerenciar os recursos genéticos animais do país (Mariante et al., 2009).

No Brasil, ações de programas de conservação vêm sendo desenvolvidas por centros de pesquisa da EMBRAPA em parceria com Universidades, Empresas de Pesquisa Estaduais e produtores (Mariane et al., 2011). O programa desenvolvido pela EMBRAPA compõe a Rede de Recursos Genéticos Animais-Rede Animal, que integra a Plataforma de Recursos Genéticos e foi criado com o objetivo de desenvolver e monitorar núcleos de conservação de raças ameaçadas de extinção, assim como realizar a caracterização fenotípica destes núcleos, a criopreservação de material genético, a caracterização genética de raças envolvidas no programa, e a conscientização da sociedade sobre a importância da conservação dos recursos genéticos animais (Mariane et al., 2009). As raças de ovinos incluídas no programa de conservação Brasileiro distribuídas por todas as regiões do país são: Barriga Negra, Bergamacia Brasileira, Somalis Brasileira, Rabo Largo, Morada Nova e Santa Inês (Mariane et al., 2009).

A conservação dos recursos genéticos animais é hoje realizada por meio de: (1) conservação *in situ*, por núcleos de conservação onde os animais são mantidos em seu habitat de origem e (2) conservação *ex situ*, in vivo, onde os animais são mantidos fora de seu habitat de origem, ou *in vitro*, baseado no armazenamento de material genético como sêmen, embriões e ovócitos em Bancos de Germoplasma (Mariane et al., 2011).

A caracterização por meio de marcadores moleculares também tem sido uma importante ferramenta para uso em programas de conservações para a quantificação da diversidade genética dentro e entre raças da mesma espécie, com a finalidade de identificar padrões de estrutura genética e estimar a distância e influências genéticas das raças comerciais em raças localmente adaptadas (McManus et al., 2010). Atualmente, a caracterização molecular tem sido usada auxiliando o manejo de rebanhos mantidos em Núcleos de Conservação e para seleção de doadores para conservação em bancos de germoplasma (McManus et al., 2010). Portanto, a importância da caracterização molecular para manter uma ampla diversidade genética de populações de animais de produção é fundamental não só ao desenvolvimento rural, mas para assegurar alimentação mundial das próximas gerações, permitindo a seleção de ações ou desenvolvimento de novas raças em resposta às grandes mudanças futuras que podem ameaçar as necessidades nutricionais humanas.

2.3 Breve estado da arte do programa de melhoramento de ovinos no Brasil.

O melhoramento genético em ovinos é uma ferramenta eficiente capaz de atender aos sistemas de produção específicos, a demanda de mercado e expectativas do consumidor. Apesar da ovinocultura ter sido introduzida no Brasil desde a época colonial, só nas últimas décadas tem se observado a ação do melhoramento genético nos rebanhos de ovinos do país por ação das associações e produtores individuais que realizam atividades de seleção de animais superiores baseados na morfologia para o aprimoramento das características de produção de interesse. No entanto, programas de melhoramento de ovinos a longo prazo ainda são escassos no Brasil e isso se dá principalmente, a desorganização da cadeia produtiva e desinteresse por parte dos produtores (McManus et al., 2010; Hermuche et al., 2013a).

A Associação Rio-Grandense de Criadores de Ovinos, atualmente denominada Associação Brasileira de Criadores de Ovinos (ARCO) fundada em 1942, iniciou as primeiras avaliações objetivas para seleção de ovinos visando a melhoria na produção e qualidade da lã (Ojeda, 1999). Os programas de melhoramento em ovinos tiveram seu início com o PROMOVI (Programa de Melhoramento Genético em Ovinos), visando a melhoria na produção de carne e lã, entre 1977 a 1995 no Rio grande de Sul, se estendendo depois para outros estados (McManus et al., 2010). Atualmente, os programas de melhoramento disponíveis no Brasil são geridos por órgãos públicos, privados e associações de criadores. A maioria desses programas são centralizados na melhoria das raças deslanadas Dorper e Santa Inês representados por programas especializados principalmente na ovinocultura de corte.

Os principais programas no Brasil são: PMGSI/USP (Programa de Melhoramento Genético da raça Santa Inês) iniciado em 2004 pela USP em parceria com a ASCCO (Associação Sergipana de Criadores de Ovinos e Caprinos), SANTAGEN (Programa de Melhoramento Genético de Ovinos Santa Inês) desenvolvida pela UFPI, PROMOSI (Programa de Melhoramento Genético de Ovinos da Raça Santa Inês) gerido pela EMEPA-PB (Empresa Estadual de Pesquisa Agropecuária da Paraíba), GENECOC- (Programa de Melhoramento Genético de Caprinos e Ovinos de Corte) lançado em 2003 pela EMBRAPA caprinos e ovinos, onde destaca-se o núcleo participativo da raça de ovinos Morada Nova; e programas desenvolvidos por empresas privadas como a OVINOPLUS (Programa de Avaliação Genética de Ovinos) pela empresa *Alta Genetics* e OviGol®/AbacusBio desenvolvido pela empresa Aries Reprodução e Melhoramento Genético Ovino – Ltda (Azevedo et., al., 2005; Lôbo, 2007; Lobo, 2006; Sousa et al., 2007).

Todos esses programas encontram dificuldades para manutenção e expansão principalmente devido ao pouco interesse e participação dos produtores, em alguns casos por desconhecimento da importância do melhoramento genético para o aumento da produtividade (Lobo, 2006).

2.4 Prolificidade e taxa de ovulação em ovinos.

A prolificidade, no que se refere a quantidade de cordeiros nascidos por ovelhas paridas, é um parâmetro importante para determinar a eficiência produtiva e reprodutiva de um rebanho (Notter, 2008). A prolificidade economicamente satisfatória para um sistema de produção semi-intensivo ou intensivo está relacionada com uma alta taxa de concepção gemelar que é resultado principalmente de uma alta taxa de ovulação materna e baixas taxas de mortalidade fetal e embrionária. A condição genética materna e/ou paterna é o fator com maior influência sobre a prolificidade em ovinos (Davis, 2005). Outros fatores que podem influenciar diretamente essa característica são: a idade da ovelha, estação do ano, nível nutricional e hormonal (Debus et al., 2012; Lewis et al., 1996; Notter, 2000). Porém, em um sistema de produção extensivo, onde os animais são criados soltos e com pouco cuidado, as taxas de concepção mais baixas são geralmente aceitáveis, assim como em ambientes extremos (Walkden-Brown; et al., 2008), já que partos múltiplos podem aumentar a taxa de mortalidade devido à seca e escassez de alimento como ocorre na região Nordeste do Brasil (da Silva et al., 2014; Paim et al., 2013).

Os índices de prolificidade nos rebanhos de ovinos do Brasil são considerados baixos, caracterizando um grande obstáculo para o desenvolvimento de uma ovinocultura rentável e de alta produção (Paim et al., 2011). O índice de prolificidade é um parâmetro reprodutivo que pode ser manejado para aumentar a escala de produção, resultando em níveis mais elevados de produtividade com redução de custos e, conseqüentemente, aumento das margens de lucro, devido à maior diluição dos custos fixos totais com um maior número de cordeiros por ovelha e maior disponibilidade de produto para venda (Lôbo et al., 2011).

A prolificidade em ovinos é uma característica limitada pelo sexo e idade reprodutiva da fêmea e dependente principalmente da taxa de ovulação. Essa característica é controlada por poucos genes, o que facilita o seu uso em programas de seleção assistida por marcadores moleculares (Davis, 2005). O emprego de metodologias de seleção assistida pela

genotipagem com a introgressão de alelos em genes de efeito principal pode produzir grandes e rápidas modificações nos índices de produção.

O processo de ovulação em mamíferos é resultado do crescimento e maturação dos folículos ovarianos, onde a cada ciclo estral, muitos folículos são recrutados (folículos antrais pequenos), no entanto poucos chegam ao estágio pré-ovulatório e são ovulados, o restante sofre atresia. O desenvolvimento e maturação destes folículos envolve um complexo de sinais endócrinos (FSH, LH, Progesterona e Estradiol), dependentes do eixo hipotálamico-hipofisário-gonadal, além de sinais parácrinos determinados pela ação de fatores de crescimento (GDFs e BMPs) que influenciam o desenvolvimento e a disposição das células da granulosa, esteroidogênese, apoptose e expansão das células do cumulus (McNatty et al., 2001; Matzuk et al., 2002; Gilchrist et al., 2008).

2.5 Genes associados a taxa de ovulação.

Os genes que controlam a expressão fenotípica da prolificidade (associados à taxa de ovulação) pertencem, principalmente, à superfamília do fator transformador de crescimento beta (*TGF-β*), como é o caso dos genes da proteína morfogenética do osso (BMPs- *Bone morphogenetic proteins*), Fator de Diferenciação e Crescimento-9 (*GDF9- Growth and differentiation factor*) e seus receptores (Davis, 2005). A família dos TGF-β, representada por 40 membros, é dividida nas subfamílias ativina/inibina, e são produzidas pelo próprio ovócito (Fortune, 2003).

Os *BMPs* e seus receptores, proteínas que atuam como homodímeros encontrados principalmente nas células da granulosa dos folículos, tem ação de regulação, proliferação e diferenciação celular, desde o estágio de folículo primário até o de folículo antral (Juengel & McNatty, 2005). O *GDF9* também é uma proteína homodimérica, e é responsável pela proliferação das células da granulosa e da teca, tanto nos folículos pré-antrais como nos antrais (Bodensteiner et al., 1999), enquanto a expressão do *BMP15* se inicia nos ovócitos dos folículos primários (Galloway et al., 2000).

Estudos genéticos identificando mutações nos genes relacionados à taxa de ovulação em mamíferos demonstraram que os ovinos têm sido um excelente modelo para elucidar a importância destes genes para a função reprodutiva nas fêmeas (Otsuka et al., 2011). Na Tabela 1 encontra-se um resumo dos estudos nos principais genes relacionados à taxa de ovulação localizados nos cromossomos X (*BMP15*), 5 (*GDF9*) e 6 (*BMPRI-b*).

Os fenótipos diferem de acordo com as mutações descritas nos três genes de efeito principal: *GDF9*, *BMP15* e *BMPRI-b* (Tabela 1). Ovelhas normalmente, ovulam de 1 a 2 folículos por ciclo (Davis et al., 2008), entretanto quando observamos mutações em genes associados à taxa de ovulação, nota-se diferentes fenótipos de ovulações únicas, múltiplas ou até infertilidade (Otsuka et al., 2011). O fenótipo de esterilidade foi observado por diferentes mutações no gene *BMP15* nas raças: Galway, (Galloway et al., 2000), Belclare, Cambridge (Hanrahan et al., 2004); Lacaune (Bodin et al., 2007); Aragonesa (Martinez-Royo et al., 2008); e no gene *GDF9* nas raças Belclare, Cambridge (Hanrahan et al., 2004) Icelandic Thoka (Nicol et al., 2009), e Ile de France (Souza et al., 2014) (Tabela 1). Estas mutações levam a inativação das respectivas proteínas dos genes *GDF9* e *BMP15* resultando na não formação dos folículos secundários, levando assim a infertilidade em ovelhas homozigotas para os alelos mutantes, enquanto aquelas em heterozigose resultam em um aumento na taxa de ovulação (Figura 1) (Otsuka et al., 2011).

Tabela 1- Mutações encontradas em genes associados à taxa de ovulação relacionadas com fenótipos de prolificidade em ovinos.

GENE	Alelo	Mudança no AA	Mutação/Raças	Fenótipo em Homozigotos	Referência
<i>BMP15</i> (<i>FecX</i>)	<i>FecX^I</i>	V299D	Romney Inverdale)	Estétil	Galloway et al., 2000
	<i>FecX^H</i>	Q291ter	Romney (Hanna)	Estétil	Hanrahan et al., 2004
	<i>FecX^B</i>	S367I	Belclare	Estétil	Hanrahan et al., 2004
	<i>FecX^G</i>	Q239ter	Galway/Belclare-Cambridge	Estétil	Galloway et al., 2000
	<i>FecX^L</i>	C321Y	Lacaune	Estétil	Bodin et al., 2007
	<i>FecX^R</i>	W154S159	Aragonesa	Estétil	Matinez-Royo et al., 2008
	<i>FecX^O</i>	A337H	Olkuska	Hiperprolífico	Demars et al., 2013
<i>BMPRI-b</i> (<i>FecB</i>)	<i>FecX^{Gr}</i>	T317I	Grivette	Hiperprolífico	Demars et al., 2013
	<i>FecB^B</i>	Q249R	Booroola Merino, Garole, Hu High	Hiperprolífico	Wilson et al.2001
<i>GDF9</i> (<i>FecG</i>)	<i>FecG^H</i>	S395F	Fertility/Belclare-Cambridge	Estétil	Hanrahan et al., 2004
	<i>FecG^{TT}</i>	S427R	Icelandic Thoka	Estétil	Nicol et al., 2009
	<i>FecG^E</i>	F345C	Santa Inês	Hiperprolífico	Silva et al., 2011
	<i>FecG^V</i>	A315C	Vacaria/ Ile de France	Estétil	Souza et al., 2014
	<i>FecG^{Han}</i>	G243H	Han - Dorset	Hiperprolífico	Chu et al., 2011
	<i>FecG^F</i>	V371M	Finnsheep, Norwegian White	Hiperprolífico	Mullen & Hanrahan, 2014; Vage et al., 2013

O fenótipo de ovelhas hiperprolíficas ocorre pelo aumento na taxa de ovulação tanto em ovelhas heterozigotas quanto em homozigotas em diferentes mutações encontrados

nos genes *BMPBR1-b* (Wilson et al., 2001), *BMP15* (Demars et al., 2013) e *GDF9* (Chu et al., 2011; Mullen & Hanrahan, 2014; Silva et al., 2011; Vage et al., 2013). As altas taxas ovulatórias são resultado do aumento da sensibilidade ao FSH, o qual atuaria promovendo o desenvolvimento folicular e permitindo a ovulação precoce nos pequenos folículos em crescimento (Jansson 2014; Shimasaki et al., 2004) (Figura 1). Portanto, diferentes estudos evidenciam o papel dos genes *GDF9*, *BMP15* e *BMPRI-b* no controle da foliculogênese, na taxa de ovulação e consequentemente na determinação da fertilidade (Otsuka et al., 2011).

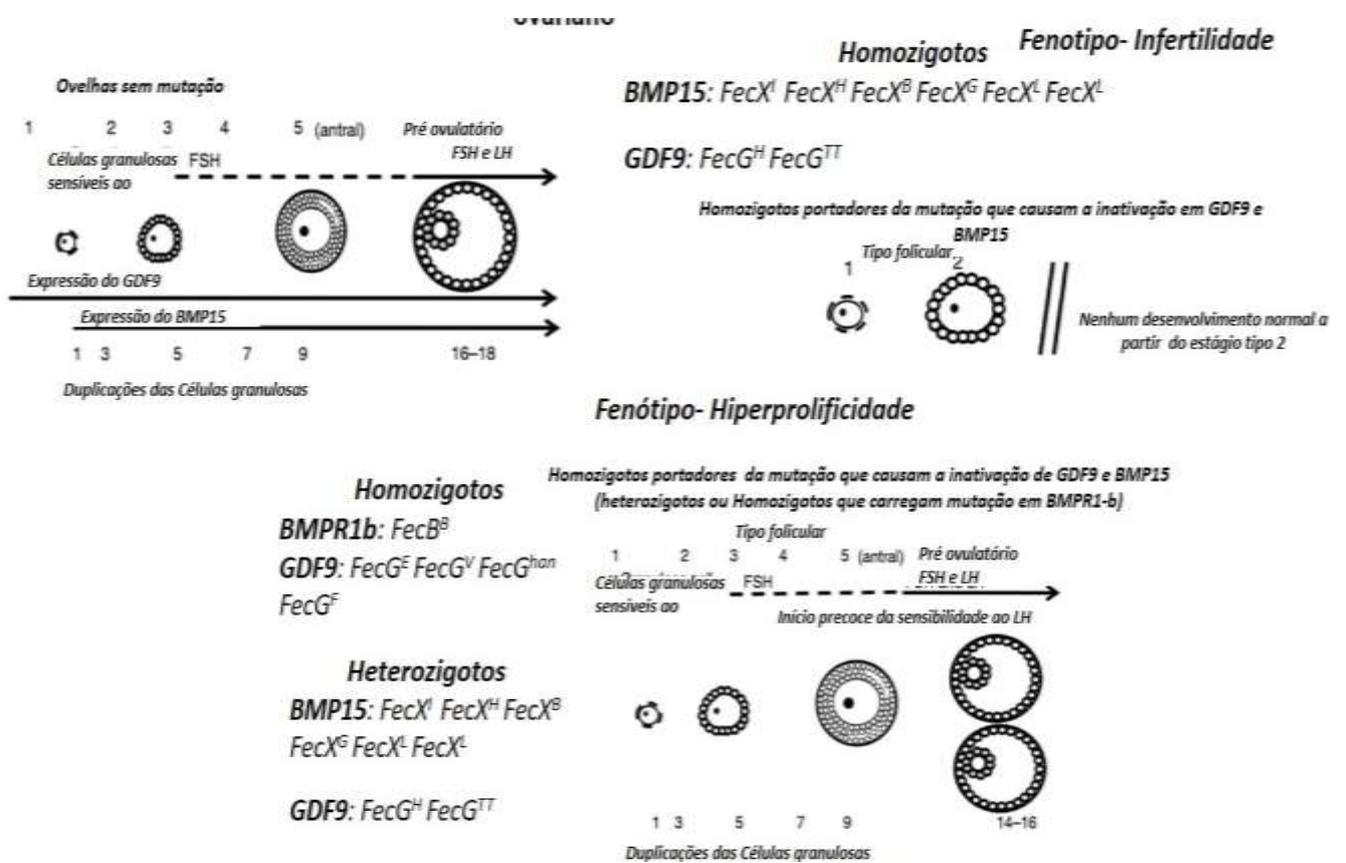


Figura 1- Diferentes alelos nos genes *GDF9*, *BMP15* e *BMPRI-b* envolvidos na regulação do desenvolvimento folicular ovariano resultando em fenótipos de hiperproliféricidade e infertilidade em diferentes raças ovinas. Fonte: (adaptada- Juengel et al., 2013).

Outras ovelhas prolíficas têm apresentado mutações em outros genes de efeito principal. Estes outros genes incluem o *BG4ALNT2*, com alta expressão gênica no ovário (Drouilhet et al., 2013) em linhagens de ovelhas Laucane ($FecL^L$), e genes ainda desconhecidos, localizados no cromossomo X, em linhagens Woodlands ($FecX^{2W}$ e $FecX^{2M}$) (Davis et al., 2008; Feary et al., 2007).

A existência de diferentes linhagens de ovinos com alta prolificidade e diferentes mutações em genes de efeito principal, levanta a hipótese que outras raças prolíficas possam apresentar as mesmas ou ainda desconhecidas mutações em genes relacionados à taxa de ovulação. Essa constatação cria a possibilidade de identificação de linhagens com polimorfismos associados a essa característica, o que abre novos caminhos para pesquisas e a utilização efetiva deste conhecimento, caracterizando geneticamente raças associadas à característica de prolificidade.

2.6 Impacto das novas tecnologias genômicas em características controladas por poucos genes.

A década de 80 foi um marco para os primeiros estudos de identificação, caracterização e utilização de marcadores moleculares em recursos genéticos e na geração de ferramentas para o melhoramento animal (Caetano, 2009). Nas últimas décadas, as inovações tecnológicas da genética molecular trouxeram novas metodologias de alto desempenho e acurácia, as quais permitem acessar e manipular o genoma, facilitando a genotipagem por marcadores moleculares, em especial os SNPs (*Single Nucleotide Polymorphism*- Polimorfismo de base única), a custo reduzido.

Os marcadores SNPs são variações de bases única sendo elas transições (trocas de purina por purina-ex.: A/G ou de uma pirimidina por pirimidina C/T) ou transversões (troca de uma purina por uma pirimidina, ou vice-versa). Estes marcadores preferencialmente bi alélicos estão distribuídos de forma homogênea por todo o genoma em eucariotos (Sachidanandam et al., 2001) e podem estar localizados em regiões de exon, com função regulatória, ou íntrons e espaços intergênicos.

O uso desses marcadores em pesquisas genéticas, como polimorfismos associados a características de produção, foi limitado por muitos anos devido aos custos e tecnologias primárias. O surgimento das novas tecnologias para caracterização, prospecção e genotipagem de marcadores SNP por um menor custo e de forma massal foi uma grande revolução para as pesquisas genéticas no mundo (Liu, 2009; The Bovine Hapmap Consortium, 2009).

2.6.1 Sequenciamento de nova geração

As tecnologias de sequenciamento de nova geração (NGS), representadas pelas plataformas de sequenciamento da 454 FLX da Roche (Margulies et al., 2005), Solexa da *Illumina* (Bennett, 2004), *SOLiD System* da *Applied Biosystems* (Valouev et al., 2008), Ion Torrent-PGM da Life Technologies e HeliScope True Single Molecule Sequencing (Helicos) (Thompson & Steinmann, 2010) e SMRT (*Single Molecule Real-Time*) da PacBio (Eid et al., 2009); foram capazes de promover o sequenciamento de DNA em plataformas, gerando informação de milhões de pares de bases (pb) em uma única corrida, números significativamente maiores quando comparados ao sequenciamento de Sanger (Ansorge, 2009). Obviamente, as novas tecnologias trouxeram grandes avanços para a genômica, reduzindo tempo e custos por base de sequenciamento, além do mais sem a necessidade de investimentos em equipamentos por grupos de pesquisas individuais (Ansorge, 2009; Metzker, 2010). Existem diversas empresas realizando prestação de serviços na área de sequenciamento de nova geração nacional e internacionalmente e com a vantagem de que os laboratórios contratados já serem especializados, o que diminui a taxa de erro e o tempo entre a contratação do serviço e obtenção dos dados (Shendure & Ji, 2008).

Espécies importantes para a pecuária já tiveram todo genoma sequenciado, como é o caso do genoma do bovino (The Bovine Hapmap Consortium, 2009), da galinha (Hillier et al., 2004), cavalo (Wade et al., 2009), ovelha, (Archibald et al., 2010a; Jiang et al., 2014) suíno (Archibald et al., 2010b), abelha (Weinstock et al., 2006), e bicho da seda (The International Silkworm Genome Consortium, 2008). Após o sequenciamento do genoma completo das espécies outros esforços vêm sendo feitos com a finalidade de se obter sequências de outros indivíduos da mesma espécie, e/ou diferentes raças, permitindo assim aumentar o conhecimento das espécies, identificação e caracterização das raças (Stothard et al., 2011; Heaton et al., 2013); Xing et al., 2016).

O grande desafio para o sequenciamento total de grandes eucariotos, por segmentos de leitura curtas (*reads*) (< 100 pb) geradas por NGS, é a montagem dos genomas completos, que exige o emprego de ferramentas de bioinformática (Li et al., 2009a; Mardis, 2008; Shen et al., 2010). O sequenciamento completo do genoma do panda (Li et al., 2010), serviu como demonstração da possibilidade da utilização de sequências curtas geradas por apenas NGS (sequenciamento gerado pelo *Illumina Genoma Analyser*) para obtenção de genomas completos.

Após o sequenciamento completo do genoma ovino, novos esforços foram feitos com o objetivo de se encontrar marcadores SNPs (Jiang et al., 2014). As estimativas de distância dos SNPs podem determinar a densidade de SNPs necessários para construção de painéis associados a uma característica de interesse (Allen et al., 2010; Bolormaa et al., 2011a; Ilic et al., 2011), como, por exemplo, a taxa de ovulação em ovinos.

2.6.2 Painéis de genotipagem de SNPs de baixa densidade

Os chips para a genotipagem de SNPs de alta e média densidade, de centenas a milhares de marcadores, contribuíram para o desenvolvimento de plataformas de genotipagens de grande rapidez e automação com um grande número de amostras em um único ensaio (Liu, 2009). Essa nova tecnologia de genotipagem em larga escala permitiu a geração de metodologias para avaliação genética e seleção com base no Valor Genômico (*Genomic Estimated Breeding Value- GEBV*). Grandes avanços foram realizados em pesquisas genômicas em espécies animais importantes para a pecuária durante os últimos anos (Allen et al., 2010; Spencer & Palmarini, 2012; Wiedmann et al., 2008). Esses esforços disponibilizaram chips de genotipagem de alta (>500k) (Anderson et al., 2010; Kijas et al., 2014; Matukumalli et al., 2009), média (>50k) (Bolormaa et al., 2011b; Kijas et al., 2012; Hayes et al., 2009; Finlay et al., 2012; Snelling et al., 2010) e baixa densidade (>100) (Beynon et al., 2015; Boichard et al., 2012) nas espécies bovina e ovina.

O Consórcio Internacional do Genoma Ovino (*International Sheep Genomics Consortium, ISGC*, (<http://www.sheepmap.org/>) validou um painel de 60.000 marcadores SNP (*Illumina ovineSNP50Beadchip*) em mais de 2800 animais provenientes de mais de 70 raças de ovinos (Kijas et al., 2012), e atualmente esse painel se encontra comercialmente disponível para toda comunidade científica.

A aplicação de testes com baixo número de marcadores em um grande número de amostras foi desenvolvida primeiramente para ensaios relacionados a doenças. Como exemplo pode-se citar as metodologias *Taqman* (Heid et al., 1996) ou de minissequenciamento.

Os painéis de baixa densidade foram usados também nos processos de rastreabilidade de amostras (Heaton et al., 2005), e para detectar polimorfismos em genes associados a características de produção de interesse comercial (Chessa et al., 2007). Essas aplicações requerem a genotipagem de algumas dezenas de marcadores em grandes números

de amostras, fazendo com que as metodologias *Taqman* e minissequenciamento não sejam mais indicadas para atender a essas demandas. Por outro lado, as plataformas baseadas em espectrometria de massa MALDI-TOF (*Matrix-assisted laser desorption/ionization – time of flight mass spectrometry*) (Heaton et al., 2005) e microarranjos (Chessa et al., 2007; Kaminski et al., 2005) são as melhores tecnologias aplicadas para esses estudos. O sistema *MassARRAY iPLEX™* é recomendado para o uso em trabalhos com mapeamento finos e validação de GWAS em estudos de ligação, bem como os testes de rotina genética de painéis de SNPs de interesses (Oeth et al., 2007).

O ensaio *Illumina® SNP Genotyping GoldenGate* em conjunto com a tecnologia *VeraCode®* (Lin et al., 2009), um método baseado na detecção a partir de fluorescência, permite um grau elevado de multiplexação de locos a partir de 48 SNPs por ensaio. A *Plataforma Fluidigm (Fluidigm Dynamic Array™ Integrated Fluidic Circuits –IFCs-* <https://www.fluidigm.com>) e *Kasp™ (Kompetitive Allele Specific PCR Genotyping System-* <http://www.lgcgroup.com/products/kasp-genotyping-chemistry>) baseada em PCR alelo-específico, permite analisar uma densidade baixa a média de SNPs em grande número de amostras (He et al., 2014; Ilic et al., 2011).

Como exemplo, o desenvolvimento de um painel de 48 SNP e sua validação pela plataforma *Fluidigm* foi usado para caracterização de germoplasma de cacau, apresentando resultados de genotipagem de alta qualidade (Ilic et al., 2012). Outro trabalho com a mesma plataforma permitiu o desenvolvimento de um painel de 96 SNPs em organismos não modelos, como o salmão, permitindo atender mesmo animais com genoma ainda não inteiramente sequenciado (Narum et al., 2009).

2.6.3 Estudos de associações de genômica ampla (GWAS).

Desequilíbrio de ligação (DL) é a medida estatística da associação entre alelos de diferentes locos, que não se dá ao acaso, e é afetado por forças evolutivas, como: recombinação, seleção, migração e deriva genética (Slatkin, 2008). O DL é correlacionado com taxa de recombinação e quanto maior a taxa de recombinação mais rápida será a aproximação do equilíbrio entre os dois locos (Slatkin, 2008). Assim, a cada nova geração, a recombinação tende a criar uma situação de equilíbrio entre os locos, e o tamanho de um dado segmento

cromossômico que contém dois locos quaisquer sem recombinação diminui, reduzindo o DL (Mcrae et al., 2002).

A disponibilidade e genotipagem de um grande número de marcadores SNPs distribuídos por todo o genoma de um indivíduo aumenta a probabilidade de encontrar marcadores em desequilíbrio de ligação (DL) com QTLs ou genes de interesse (Dekkers, 2004, 2007). Assim, o uso eficiente de marcadores moleculares para auxiliar o melhoramento genético requer um maior número possível de marcadores, para que se aumente a possibilidade de encontrar um marcador em DL com um determinado QTL ou gene de interesse (Dekkers, 2004, 2007). A associação encontrada entre o marcador e o gene e/ou QTL tem que ser validada em diferentes populações e a cada geração de melhoramento para que se garanta que a mesma não seja perdida pelo processo natural de recombinação (Mackay, 2001).

As ferramentas de genotipagem de SNPs em larga escala disponíveis atualmente possibilitaram o desenvolvimento da metodologia denominada estudos de associação genômica ampla (*GWAS- Genome Wide Association Study*), os quais já foram realizados utilizando características controladas por poucos genes em algumas espécies de animais de produção como bovinos (Charlier et al., 2008; Tosser-Klopp et al., 2014), equinos (Dupuis et al., 2011) e ovinos (Becker et al., 2010). A GWAS é baseada na associação entre marcadores espalhados por todo o genoma genotipados num grupo de indivíduos e uma característica de interesse produtiva ou uma doença

3 REFERÊNCIAS BIBLIOGRÁFICAS

- ALLEN, A. R. et al. Compilation of a panel of informative single nucleotide polymorphisms for bovine identification in the Northern Irish cattle population. **BMC Genetics**, v. 11, p. 5, 2010.
- ANDERSON, R. et al. Development of a High Density (600K) Illumina Ovine SNP Chip and its Use to Fine Map the Yellow Fat Locus. v. 2, p. 2010, 2010.
- ANSORGE, W. J. Next-generation DNA sequencing techniques. **New biotechnology**, v. 25, n. 4, p. 195–203, abr. 2009.
- ARCHIBALD, A. L. et al. The sheep genome reference sequence: a work in progress. **Animal genetics**, v. 41, n. 5, p. 449–53, 2010a.
- ARCHIBALD, A. L. et al. Pig genome sequence--analysis and publication strategy. **BMC genomics**, v. 11, n. July, p. 438, 2010b.
- AZEVEDO, D.; SILVA, M. M. R. .; D. C.; PESSOA, R. S.; MOURA JÚNIOR, P. N. **Programa de Melhoramento Genético de Ovinos Santa Inês : SANTAGEN**. 2005 Disponível em: <[Http://www.ufpi.br/subsiteFiles/caprioavis/arquivos/files/Artigo 9.pdf](http://www.ufpi.br/subsiteFiles/caprioavis/arquivos/files/Artigo%209.pdf)>.
- BECKER, D. et al. Microphthalmia in Texel sheep is associated with a missense mutation in the paired-like homeodomain 3 (PITX3) gene. **PloS one**, v. 5, n. 1, p. e8689, 2010.
- BENNETT, S. Solexa Ltd. **Pharmacogenomics**, v. 5, n. 4, p. 433–438, 2004.
- BEYNON, S. E. et al. Population structure and history of the Welsh sheep breeds determined by whole genome genotyping. **BMC Genetics**, v. 16, n. 1, p. 65, 2015.
- BODENSTEINER, K. J. et al. Molecular cloning of the ovine Growth/Differentiation factor-9 gene and expression of growth/differentiation factor-9 in ovine and bovine ovaries. **Biology of reproduction**, v. 60, n. 2, p. 381–386, 1999.
- BODIN, L. et al. A novel mutation in the bone morphogenetic protein 15 gene causing defective protein secretion is associated with both increased ovulation rate and sterility in Lacaune sheep. **Endocrinology**, v. 148, n. 1, p. 393–400, 2007.
- BOICHARD, D. et al. Design of a bovine low-density SNP array optimized for imputation. **PloS one**, v. 7, n. 3, p. e34130, 2012.
- BOLORMAA, S. et al. Genome-wide association studies for feedlot and growth traits in cattle. **Journal of Animal Science**, v. 89, n. 6, p. 1684–1697, 2011a.
- BOLORMAA, S. et al. A genome-wide association study of meat and carcass traits in Australian cattle. **Journal of Animal Science**, v. 89, n. 8, p. 2297–2309, 2011b.
- CAETANO, A. R. Marcadores SNP: Conceitos básicos, aplicações no manejo e no melhoramento animal e perspectivas para o futuro. **Revista Brasileira de Zootecnia**, v. 38, n. SUPPL. 1, p. 64–71, 2009.
- CHARLIER, C. et al. Highly effective SNP-based association mapping and management of recessive defects in livestock. **Nature genetics**, v. 40, n. 4, p. 449–454, 2008.
- CHESSA, S. et al. Development of a single nucleotide polymorphism genotyping microarray

- platform for the identification of bovine milk protein genetic polymorphisms. **Journal of dairy science**, v. 90, n. 1, p. 451–64, 2007.
- CHU, M. X. et al. GDF9 as a candidate gene for prolificacy of Small Tail Han sheep. **Molecular biology reports**, v. 38, n. 8, p. 5199–204, 2011.
- DA SILVA, A. F. et al. Production traits in F 1 and F 2 crosses with naturalized hair breed Santa Inês ewes. p. 3–8, 2014.
- DAVIS, G. H. Major genes affecting ovulation rate in sheep. **Genetics, selection, evolution : GSE**, v. 37 Suppl 1, p. S11–S23, 2005.
- DAVIS, G. H. et al. Hyper-Prolific Ewes Carrying Copies of Three Major Genes: A Model for Studying Genes Controlling Ovulation Rate. **Biology of Reproduction**, v. 78, n. Suppl 1, p. 110, 1 maio 2008.
- DEBUS, N. et al. Maternal periconceptional undernutrition in Merinos d'Arles sheep: 1. Effects on pregnancy and reproduction results of dams and offspring growth performances. **Theriogenology**, v. 77, n. 7, p. 1453–1465, 7 abr. 2012.
- DEKKERS, J. C. M. Commercial application of marker- and gene-assisted selection in livestock: strategies and lessons. **Journal of Animal Science**, v. 82, n. E. Suppl., p. E313–E328, 2004.
- DEKKERS, J. C. M. Marker-assisted selection for commercial crossbred performance. **Journal of animal science**, v. 85, n. 9, p. 2104–2114, set. 2007.
- DEMARS, J. et al. Genome-Wide Association Studies Identify Two Novel BMP15 Mutations Responsible for an Atypical Hyperprolificacy Phenotype in Sheep. **PLoS Genetics**, v. 9, n. 4, p. e1003482, 2013.
- DROUILHET, L. et al. The Highly Prolific Phenotype of Lacaune Sheep Is Associated with an Ectopic Expression of the B4GALNT2 Gene within the Ovary. **PLoS Genetics**, v. 9, n. 9, p. e1003809, 2013.
- DUPUIS, M.-C. et al. Results of a haplotype-based GWAS for recurrent laryngeal neuropathy in the horse. **Mammalian Genome**, v. 22, n. 9-10, p. 613–620, 2011.
- EID, J. et al. Real-time DNA sequencing from single polymerase molecules. **Science (New York, N.Y.)**, v. 323, n. 5910, p. 133–138, 2009.
- FAO. **Estatísticas de Produção de Ovinos no Mundo (Ano-2013)**. Disponível em: <<http://faostat.fao.org/>>. Acesso em: 5 dez. 2015.
- FEARY, E. S. et al. Patterns of expression of messenger RNAs encoding GDF9, BMP15, TGFBR1, BMPR1B, and BMPR2 during follicular development and characterization of ovarian follicular populations in ewes carrying the Woodlands FecX2W mutation. **Biology of reproduction**, v. 77, n. 6, p. 990–998, dez. 2007.
- FINLAY, E. K. et al. A genome wide association scan of bovine tuberculosis susceptibility in Holstein-Friesian dairy cattle. **PloS one**, v. 7, n. 2, p. e30545, 2012.
- FORTUNE, J. E. The early stages of follicular development: Activation of primordial follicles and growth of preantral follicles. **Animal Reproduction Science**, v. 78, n. 3-4, p. 135–163, 2003.
- GALLOWAY, S. M. et al. Mutations in an oocyte-derived growth factor gene (BMP15) cause increased ovulation rate and infertility in a dosage-sensitive manner. **Nature genetics**, v. 25, n. 3, p. 279–283, 2000.
- GILCHRIST, R. B.; LANE, M.; THOMPSON, J. G. Oocyte-secreted factors: Regulators of cumulus cell function and oocyte quality. **Human Reproduction Update**, v. 14, n. 2, p. 159–177, 2008.
- HANRAHAN, J. P. Mutations in the Genes for Oocyte-Derived Growth Factors GDF9 and BMP15 Are Associated with Both Increased Ovulation Rate and Sterility in Cambridge and Belclare Sheep (*Ovis aries*). **Biology of Reproduction**, v. 70, n. 4, p. 900–909, 2004.
- HAYES, B. J. et al. A validated genome wide association study to breed cattle adapted to an environment altered by climate change. **PLoS ONE**, v. 4, n. 8, p. 1–8, 2009.
- HE, C.; HOLME, J.; ANTHONY, J. SNP genotyping: the KASP assay. **Methods in molecular**

- biology (Clifton, N.J.)**, v. 1145, p. 75–86, 2014.
- HEATON, M. P. et al. Use of bovine single nucleotide polymorphism markers to verify sample tracking in beef processing. **Journal of the American Veterinary Medical Association**, v. 226, n. 8, p. 1311–1314, 2005.
- HEATON, M. P. et al. Genetic Testing for TMEM154 Mutations Associated with Lentivirus Susceptibility in Sheep. **PLoS ONE**, v. 8, n. 2, p. e55490, 2013.
- HEID, C. A. et al. Real Time Quantitative PCR. p. 986–994, 1996.
- HERMUCHE, P. et al. Dynamics of Sheep Production in Brazil. **ISPRS International Journal of Geo-Information**, v. 2, n. 3, p. 665–679, 2013a.
- HERMUCHE, P. et al. Environmental factors that affect sheep production in Brazil. **Applied Geography**, v. 44, p. 172–181, 2013b.
- HILLIER, L. W. et al. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. **Nature**, v. 432, n. 7018, p. 695–716, 2004.
- IBGE. Produção da Pecuária Municipal. **Instituto Brasileiro de Geografia e Estatística**, v. 41, 2013.
- ILIC, K. et al. **High-Throughput SNP Genotyping and Parentage Verification for North American Beef and Dairy Cattle Using a 96-SNP Parentage , Traceback , and Animal ID panel** **Plant and Animal Genomes XIX Conference**, 2011.
- ILIC, K. et al. Cacao Tree Germplasm Characterization with 48-SNP Genotyping Panel Using Fluidigm SNPtype TM Assays and Dynamic Array TM Integrated Fluidic Circuits. p. 94080, 2012.
- JANSSON, T.; JANSSON, T. Genes involved in ovulation rate and litter size in sheep. 2014.
- JIANG, Y. et al. The sheep genome illuminates biology of the rumen and lipid metabolism. **Science (New York, N.Y.)**, v. 344, n. 6188, p. 1168–73, 2014.
- JUENGEL, J. L.; DAVIS, G. H.; MCNATTY, K. P. Using sheep lines with mutations in single genes to better understand ovarian function. **Reproduction**, v. 146, n. 4, p. R111–23, 2013.
- JUENGEL, J. L.; MCNATTY, K. P. The role of proteins of the transforming growth factor- β superfamily in the intraovarian regulation of follicular development J.L. **Human Reproduction Update**, v. 11, n. 2, p. 143–160, 2005.
- KAMIŃSKI, S. et al. MilkProtChip--a microarray of SNPs in candidate genes associated with milk protein biosynthesis--development and validation. **Journal of Applied Genetics**, v. 46, n. 1, p. 45–58, 2005.
- KIJAS, J. W. et al. Genome-Wide Analysis of the World's Sheep Breeds Reveals High Levels of Historic Mixture and Strong Recent Selection. **PLoS Biol**, v. 10, n. 2, p. e1001258, 2012.
- KIJAS, J. W. et al. Linkage disequilibrium over short physical distances measured in sheep using a high-density SNP chip. **Animal genetics**, v. 45, n. 5, p. 754–7, 2014.
- LEWIS, R. M. et al. Ewe fertility in the STAR accelerated lambing system. **Journal of animal science**, v. 74, n. 7, p. 1511–1522, jul. 1996.
- LI, H. et al. The Sequence Alignment/Map format and SAMtools. **Bioinformatics**, v. 25, n. 16, p. 2078–2079, 2009.
- LI, R. et al. The sequence and de novo assembly of the giant panda genome. **Nature**, v. 463, n. 7279, p. 311–317, 2010.
- LIN, C. H. et al. Medium- to High-Throughput SNP Genotyping Using VeraCode Microbeads. In: BUGERT, P. (Ed.). **DNA and RNA Profiling in Human Blood: Methods and Protocols**. Totowa, NJ: Humana Press, 2009. p. 129–142.
- LIU, G. E. Applications and case studies of the next-generation sequencing technologies in food, nutrition and agriculture. **Recent patents on food, nutrition & agriculture**, v. 1, n. 2, p. 75–79, 2009.
- LOBO, R. N. B. **Programa de Melhoramento de Pequenos Ruminantes no Brasil Raimundo SEMANA DA CAPRINOCULTURA E DA OVINOCULTURA BRASILEIRAS.**

- Anais...**Campo Grande- MS: Embrapa Gado de Corte; Embrapa Caprinos, Seção palestras., 2006 Disponível em: <<http://www.alice.cnptia.embrapa.br/handle/doc/528789>>
- LÔBO, R. N. B. et al. Economic values for production traits of Morada Nova meat sheep in a pasture based production system in semi-arid Brazil. **Small Ruminant Research**, v. 96, n. 2-3, p. 93–100, 2011.
- LÔBO, R. N. B.; LÔBO, M. B. Melhoramento genético como ferramenta para o crescimento e o desenvolvimento da ovinocultura de corte. **Rev Bras Reprodução Animal**, v. 31, n. Xvii, p. 247–253, 2007.
- MACKAY, T. F. C. The genetic architecture of quantitative traits. **Annual review of genetics**, v. 35, n. 1, p. 303–339, 2001.
- MARDIS, E. R. Next-generation DNA sequencing methods. **Annual review of genomics and human genetics**, v. 9, p. 387–402, 2008.
- MARGULIES, M. et al. Genome sequencing in microfabricated high-density picolitre reactors. **Nature**, v. 437, n. 7057, p. 376–80, 2005.
- MARIANTE, A S.; ALBUQUERQUE, M. S. M.; RAMOS, A F. Criopreservação de recursos genéticos animais brasileiros. **Revista Brasileira de Reprodução Animal**, v. 35, n. Xix, p. 64–68, 2011.
- MARIANTE, A. DA S. et al. Present status of the conservation of livestock genetic resources in Brazil. **Livestock Science**, v. 120, n. 3, p. 204–212, 2009.
- MARTINEZ-ROYO, A. et al. A deletion in the bone morphogenetic protein 15 gene causes sterility and increased prolificacy in Rasa Aragonesa sheep. **Anim Genet**, v. 39, n. 3, p. 294–297, 2008.
- MATUKUMALLI, L. K. et al. Development and characterization of a high density SNP genotyping assay for cattle. **PloS one**, v. 4, n. 4, p. e5350, 2009.
- MATZUK, M. M. et al. Intercellular communication in the mammalian ovary: oocytes carry the conversation. **Science (New York, N.Y.)**, v. 296, n. 5576, p. 2178–2180, 2002.
- MCMANUS, C. et al. Geographical distribution of sheep breeds in Brazil and their relationship with climatic and environmental factors as risk classification for conservation. **Brazilian Journal of Science and Technology**, v. 1, n. 1, p. 1–15, 2014.
- MCMANUS, C.; PAIVA, S. R.; ARAÚJO, R. O. DE. Genetics and breeding of sheep in Brazil. **Revista Brasileira de Zootecnia**, v. 39, p. 236–246, 2010.
- MCNATTY, K. P. et al. Genetic mutations influencing ovulation rate in sheep. **Reproduction, fertility, and development**, v. 13, n. 7-8, p. 549–555, 2001.
- MCRAE, A. F. et al. Linkage disequilibrium in domestic sheep. **Genetics**, v. 160, n. 3, p. 1113–1122, 2002.
- METZKER, M. L. Sequencing technologies - the next generation. **Nature reviews. Genetics**, v. 11, n. 1, p. 31–46, 2010.
- MULLEN, M. P.; HANRAHAN, J. P. Direct evidence on the contribution of a missense mutation in GDF9 to variation in ovulation rate of Finnsheep. **PloS one**, v. 9, n. 4, p. e95251, 2014.
- NARUM, S.; CAMPBELL, N.; YOUNG, Y. High-Throughput SNP Genotyping in Salmon and Steelhead. **Plant & Animal Genome Conference**, 2009.
- NICOL, L. et al. Homozygosity for a single base-pair mutation in the oocyte-specific GDF9 gene results in sterility in Thoka sheep. **Reproduction**, v. 138, n. 6, p. 921–933, 2009.
- NOTTER, D. R. Effects of ewe age and season of lambing on prolificacy in US Targhee, Suffolk, and Polypay sheep. **Small ruminant research: the journal of the International Goat Association**, v. 38, n. 1, p. 1–7, set. 2000.
- NOTTER, D. R. Genetic Aspects of Reproduction in Sheep. **Reproduction in Domestic Animals**, v. 43, p. 122–128, 2008.
- OETH, P. et al. iPLEXTM Assay: Increased Plexing Efficiency and Flexibility for MassARRAY[®] System Through Single Base Primer Extension with Mass-Modified Terminators. **Design**, n. 8876, p. 1–12, 2007.

- OJEDA, D. B. Participação do melhoramento genético na produção ovina. **Revista Brasileira de Reprodução Animal**, v. 23, n. 2, p. 146–149, 1999.
- OTSUKA, F.; MCTAVISH, K. J.; SHIMASAKI, S. Integral role of GDF-9 and BMP-15 in ovarian function. **Molecular Reproduction and Development**, v. 78, n. 1, p. 9–21, 2011.
- PAIM, T. DO P. et al. Performance, survivability and carcass traits of crossbred lambs from five paternal breeds with local hair breed Santa Inês ewes. **Small Ruminant Research**, v. 112, n. 1-3, p. 28–34, 2013.
- PAIM, T. P. et al. Estudo econômico da produção de cordeiros cruzados confinados abatidos em diferentes pesos. **Ciência Animal Brasileira**, v. 12, p. 48–57, 2011.
- PAIVA, S. R. et al. Genetic variability of the Brazilian hair sheep breeds. **Pesquisa Agropecuária Brasileira**, v. 40, n. 9, p. 887–893, set. 2005.
- PRIMO, A. T. **América: conquista e colonização**. Porto Alegre: Movimento, 2004.
- RODERO, A. et al. ANDALUSIAN LIVESTOCK IN THE DISCOVERY OF AMERICA PRIMITIVE ANDALUSIAN LIVESTOCK AND THEIR IMPLICATIONS IN THE DISCOVERY OF AMERICA DESCUBRIMIENTO DE AMERICA and during the America conquest and The specific study of the history of the andalusian livestock. v. 41, 1992.
- SACHIDANANDAM, R. et al. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. **Nature**, v. 409, n. 6822, p. 928–933, 2001.
- SHEN, Y. et al. A SNP discovery method to assess variant allele probability from next-generation resequencing data. **Genome Research**, v. 20, n. 2, p. 273–280, 2010.
- SHENDURE, J.; JI, H. Next-generation DNA sequencing. **Nat Biotechnol**, v. 26, n. 10, p. 1135–1145, 2008.
- SHIMASAKI, S. et al. The Bone Morphogenetic Protein System in Mammalian Reproduction. **Endocrine Reviews**, v. 25, n. 1, p. 72–101, 2004.
- SILVA, B. D. M. et al. A new polymorphism in the Growth and Differentiation Factor 9 (GDF9) gene is associated with increased ovulation rate and prolificacy in homozygous sheep. **Animal Genetics**, v. 42, n. 1, p. 89–92, 2011.
- SIMPLÍCIO, A. A.; SIMPLÍCIO, K. M. M. G. **Caprinocultura e ovinocultura de corte: desafios e oportunidades**. Revista CFMV. **Anais...** Brasília- DF: 2006
- SLATKIN, M. Linkage disequilibrium--understanding the evolutionary past and mapping the medical future. **Nature reviews. Genetics**, v. 9, n. 6, p. 477–85, 2008.
- SNELLING, W. M. et al. Genome-wide association study of growth in crossbred beef cattle. **Journal of Animal Science**, v. 88, n. 3, p. 837–848, 2010.
- SOUZA, W.H.; LEITE, P.R.M. **Ovinos de corte**. A raça Dorper. João Pessoa: EMEPA, 2000. 76 p.
- SOUZA, W. H. et al. **Programa de melhoramento genético para os ovinos deslançados do Brasil: Ovinos da raça Santa Inês** SIMPÓSIO INTERNACIONAL SOBRE CAPRINOS E OVINOS DE CORTE. **Anais...** João Pessoa: ANAIS do III SIMPÓSIO INTERNACIONAL SOBRE CAPRINOS E OVINOS DE CORTE- João Pessoa: EMEPA- PB, 2007
- SOUZA, J. C. H. DE; MELO, E. O.; FERRUGEM, J. C. M. Genética da prolificidade e seu emprego na produção ovina. n. 6, p. 163–166, 2009.
- SOUZA, C. J. H. et al. Mutation in the protease cleavage site of GDF9 increases ovulation rate and litter size in heterozygous ewes and causes infertility in homozygous ewes. **Animal genetics**, v. 45, n. 5, p. 732–9, 2014.
- SPENCER, T. E.; PALMARINI, M. Application of next generation sequencing in mammalian embryogenomics: lessons learned from endogenous betaretroviruses of sheep. **Animal reproduction science**, v. 134, n. 1-2, p. 95–103, 2012.
- STOTHARD, P. et al. Whole genome resequencing of black Angus and Holstein cattle for SNP and CNV discovery. **BMC Genomics**, v. 12, n. 1, p. 559, 2011.
- THE BOVINE HAPMAP CONSORTIUM. Genome-Wide Survey of SNP Variation Uncovers the

- Genetic Structure of Cattle Breeds. **SCIENCE**, v. 324, p. 528–532, 2009.
- THE INTERNATIONAL SILKWORM GENOME CONSORTIUM. The genome of a lepidopteran model insect, the silkworm *Bombyx mori*. **Insect Biochemistry and Molecular Biology**, v. 38, n. 12, p. 1036–1045, 2008.
- THOMPSON, J. F.; STEINMANN, K. E. Single molecule sequencing with a HeliScope genetic analysis system. **Current protocols in molecular biology / edited by Frederick M. Ausubel ...**, v. Chapter 7, p. Unit7.10, 2010.
- TOSSER-KLOPP, G. et al. Design and Characterization of a 52K SNP Chip for Goats. **PLoS ONE**, v. 9, n. 1, p. e86227, 22 jan. 2014.
- VÅGE, D. I. et al. A missense mutation in growth differentiation factor 9 (GDF9) is strongly associated with litter size in sheep. **BMC genetics**, v. 14, n. 1, p. 1, 2013.
- VALOUEV, A. et al. A high-resolution, nucleosome position map of *C. elegans* reveals a lack of universal sequence-dictated positioning. **Genome research**, v. 18, n. 7, p. 1051–63, 2008.
- VIANA, J. G. A. Panorama Geral da Ovinocultura no Mundo e no Brasil. **Revista Ovinos**, n. 12, p. 1–9, 2008.
- WADE, C. M. et al. Genome sequence, comparative analysis, and population genetics of the domestic horse. **Science (New York, N.Y.)**, v. 326, n. 5954, p. 865–7, 2009.
- WALKDEN-BROWN, S. W.; VAN DER WERF, J. H. J.; NIMBKAR, C. **Use of the FecB (Booroola) gene in sheep-breeding programs** (Gupta; V.S, Ed.) Proceedings of the Helen Newton Turner Memorial International Workshop, Australian Centre for International Agricultural Research 10–12 November, 2008. **Anais...India: Australian Centre for International Agricultural Research**, 2008 Disponível em: <[http://aciarc.gov.au/files/node/11617/ACIAR_PR133\(online\).pdf](http://aciarc.gov.au/files/node/11617/ACIAR_PR133(online).pdf)>
- WEINSTOCK, G. M. et al. Insights into social insects from the genome of the honeybee *Apis mellifera*. **Nature**, v. 443, n. October, p. 931–949, 2006.
- WIEDMANN, R. T.; SMITH, T. P. L.; NONNEMAN, D. J. SNP discovery in swine by reduced representation and high throughput pyrosequencing. **BMC genetics**, v. 9, p. 81, 2008.
- WILSON, T. et al. Highly prolific Booroola sheep have a mutation in the intracellular kinase domain of bone morphogenetic protein IB receptor (ALK-6) that is expressed in both oocytes and granulosa cells. **Biology of reproduction**, v. 64, n. 4, p. 1225–1235, 2001.
- XING, K. et al. Identification of genes for controlling swine adipose deposition by integrating transcriptome, whole-genome resequencing, and quantitative trait loci data. **Scientific Reports**, v. 6, n. September 2015, p. 23219, 2016.

CAPÍTULO 2

SELEÇÃO ASSISTIDA POR MARCADOR DE BASE ÚNICA (SNP) EM OVINOS DOS NÚCLEOS DE MELHORAMENTO GENÉTICO PARTICIPATIVO DA RAÇA MORADA NOVA. *

*Artigo publicado no periódico Small Ruminant Research

RESUMO

Ovinos da raça deslanada Morada Nova apresentam características desejáveis para a produção de cordeiros, especialmente em sistemas de produção extensivo no Nordeste do Brasil, representando um importante recurso genético para a produção de cordeiros em climas semiáridos do Brasil e em outros países. Testes de desempenho tem sido realizado anualmente com esta raça desde o ano 2008. No presente estudo, ovelhas Morada Nova de dois estados brasileiros: Ceará (140 animais) e São Paulo (112 animais); foram genotipados para o SNP (alelo *FecG^E*) associado à taxa de ovulação em ovelhas, que até o momento, somente foi encontrado em raças brasileiras de ovinos localmente adaptadas. A frequência total observada de *FecG^E* foi de 0,65, e, também, foi observado um aumento do número de heterozigotos ($\chi^2=7.274, p < 0.01$). Não foram observadas diferenças significativas das frequências do alelo *FecG^E* ($p = 0,3708$) em 139 carneiros que participavam do teste de desempenho, classificados como Elite/Superior ou Regular/Inferior nos estados do Ceará e São Paulo. Considerando que o tamanho da ninhada pode afetar positivamente a rentabilidade nas fazendas, a inclusão de informações de genotipagem de *FecG^E* para índices de seleção futuras, estimadas em dados de testes de desempenho, poderá aperfeiçoar os sistemas de produção regionais que contribuirá para o aumento na rentabilidade e ganhos observados no programa de melhoramento da raça Morada Nova.

Palavras Chaves: Recursos genéticos animais, *GDF9*, *Ovis aries*, teste de desempenho, ovinos Morada Nova, taxa de ovulação.

ABSTRACT

SINGLE MARKER ASSISTED SELECTION IN BRAZILIAN MORADA NOVA HAIR SHEEP COMMUNITY-BASED BREEDING PROGRAM.

Morada Nova hair sheep show traits desirable for lamb production especially in extensive production systems in Northeastern Brazil, representing an important genetic resource for producing lamb in semi-arid climates in Brazil and elsewhere. Performance testing has been carried out annually with this breed since 2008. In the present study, Morada Nova sheep from two Brazilian states: Ceará (140 animals) and São Paulo (112 animals) were genotyped for a SNP associated with litter size, which is almost only found in Brazilian locally adapted sheep breeds (*FecG^E*). The total observed frequency of *FecG^E* was 0.65, while an increased number of observed heterozygotes was also observed ($\chi^2=7.274$, $p < 0.01$). No significant *FecG^E* allele frequency differences were observed ($p=0.3708$) in 139 performance-tested rams classified as Elite/Superior or Regular/Inferior in the states of Ceará and São Paulo. Considering that litter size has been shown to positively affect farm profitability in medium to high input systems, we suggest that inclusion of *FecG^E* genotyping information in future selection indexes estimated with basis on performance test data, fine-tuned to regional production systems may contribute to increase profitability gains observed in the Morada Nova community-based breeding program.

Keywords: Animal Genetic Resources, *GDF9*, *Ovis aires*, performance test, Morada Nova sheep.

1 INTRODUCTION

The use of locally adapted sheep breeds for meat production has been increasing in Brazil as farmers look for animals that are productive in stressful environments, especially in the Northeast region, where soils are shallow, poor and climates range from tropical semi-humid to semi-arid, with irregularly distributed rainfall rates ranging from 250 to 700 mm/year. Purebred and crossbred hair sheep breeds with high heat tolerance (Castanheira et al., 2010) and parasite resistance (McManus et al., 2009) are frequently raised in these areas (McManus et al., 2014), and account for most of the regional lamb production.

The Morada Nova breed was originally described by Domingues (1954) and is one of the most important hair sheep used for lamb production in the aforementioned regions. The breed shows good production traits, such as high rusticity and average growth in pasture-based systems (Facó et al., 2008). Ewes are sexually precocious, highly fertile and prolific, and show good maternal ability, which added to the small average adult size contribute to the profitability of local lamb production in low input systems (Facó et al., 2008; Lôbo et al., 2011).

Recent studies in animal nutrition, reproduction, genetics and breeding, combined with restructuring of the Morada Nova Breed Association in 2008, provided the basis for the establishment of a local community-based breeding program coordinated by Embrapa Sheep and Goat Research Centre (CNPCCO). The program is responsible for overlooking production and pedigree data collection and analysis, and for centralizing performance testing to identify young superior rams for use as sires in associated flocks (Faco et al., 2009).

Profitability studies showed that prolificacy (number of lambs born per adult ewe) is a major determining factor affecting production efficiency in local systems (Paim et al., 2011). Lobo et al. (2011) and McManus et al. (2011) showed that prolificacy is directly linked to the economic viability of using Morada Nova as well as other hair sheep breeds for lamb

production in Brazil, respectively. According to Rao & Notter (2000), litter size can be easily measured and can respond to directional selection besides its low estimated heritability (Fernandes, 1992). Genetic gains of 1% in the litter size have been estimated to result in profit increments of US\$ 0.781 ewe/year (4% of the total profit), considering a pasture-based production system with Morada Nova sheep in Brazil's semi-arid region (Lobo et al., 2011).

Sheep have provided a valuable model to study ovulation rate in mammals, as several genes/mutations have been identified in parallel studies in multiple breeds (Otsuka et al., 2011; Juengel et al., 2013). Polymorphisms on ovine *GDF9* (Growth and Differentiation Factor 9) exon 2 have been shown to cause ovulation rate differences. At least three distinct mutations have been described: *FecG^H* (Hanrahan et al., 2004), *FecG^T* (Nicol et al., 2009) and *FecG^E* (Silva et al., 2011). Ewes homozygous for the *GDF9*-S77F (Hanrahan et al., 2004) and *GDF9*-S109R (Nicol et al., 2009) mutations have been shown to be infertile, while heterozygotes show increased ovulation rates.

The *GDF9 FecG^E* allele was identified in the locally adapted Brazilian Santa Ines breed (Silva et al., 2011). The mutation results in a change of a phenylalanine to a cysteine in the mature peptide, and causes an increase in ovulation rate only in homozygotes. The aim of this work was to estimate *FecG^E* allele frequencies in Morada Nova flocks in Brazil and to elaborate a strategy and to estimate potential results from implementing marker assisted selection for this SNP as a regular procedure in the breed's community-based breeding program.

2 MATERIAL AND METHODS

DNA was extracted from blood from 252 Morada Nova sheep with a protocol modified from Miller et al.(1988). Samples included 122 young rams from four distinct Performance Tests (PT) (between 2008 - 2010) originated from 20 production farms and 18 animals from Embrapa Sheep and Goats (Sobral, CE, Brazil) genetic conservation nucleus. These animals were all from Ceará State (Northeast region), Brazil. Additional samples from 112 sheep derived from five farms in São Paulo State (Southeast region) were also used (17 of which were derived from a single PT in São Paulo State). Samples were derived from major rams for each farm included in the study to reduce within-farm relationships when pedigree data was not available and avoid sampling of animals with common grandparents on farms where pedigree data was available.

Performance testing of young rams was carried out according to Facó et al.,(2009). Rams received a final score based on daily weight gain (DWG), rib eye area (REAp = $REA / final\ weight^{0.75}$) and scrotal circumference (SCp = $SC / final\ weight^{0.75}$) weighted by metabolic weight, as well as fat thickness (FT) and visual scores (VS) (Facó et al.,2009) where: Final Index (FI) = $0.40(DWG) + 0.15(REAp) + 0.10(SCp) + 0.10(FT) + 0.25(VS)$. Animals were then classified as Elite (FI > mean + 1.0 standard deviation (SD)), Superior (FI between mean and 1 SD), Regular (FI between mean and -1 SD) and Inferior (FI < mean - 1SD). The number of observed *FecG^E* alleles (*FecG^E/FecG^E* (1), *FecG^E/FecG⁺* (2) and *FecG⁺/FecG⁺* (3)) in rams from the Ceará PT was regressed on final category (Elite/Superior (1) or Regular/Inferior (0)) in a logistic regression using SAS v. 9.3 (SAS Institute, Cary, North Carolina). The model is stated in terms of the probability that Final Category (Y) = 1 (Elite/Superior), referred to as \hat{p} . The probability that Y is 0 (Regular/Inferior) is

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right) = \beta_0 + \beta_1 X$$

where X is the number of *FecG^E* alleles; ln is the natural log; β_0 and β_1 are the regression coefficients and the odds ratios were calculated as $Odds = e^{\beta_0 + \beta_1 X}$.

The GDF9*FecG^E* mutation was genotyped by direct Sanger sequencing of PCR-amplified fragments. Two pairs of primers were used: one pair for amplifying a larger fragment of ~900bp (5'-GAGAAAAGGGACAGAAGC-3' and 5'-ACGACAGGTACTTAGT-3', from Silva et al. (2011) and one internal pair for amplifying a smaller fragment of ~400 bp (5'-CCTCCACCCTAAAAGGAAGC-3' and 5'-GGTCTTGGCACTGAGGAGTC-3'). PCR was carried out with annealing temperature of 58°C for 30 cycles, in a final volume of 10 ul containing 6 ng genomic DNA, 1X *QUIAGEN Multiplex PCR Master Mix* (Qiagen Inc., Valencia, CA, USA), 0.5x *Q-Solution*, 0.1µM of each primer and RNase-Free Water to complete the final reaction volume. PCR products were purified with an EXOSAP-IT and used for sequencing, following *BigDye terminator v.3* (*Applied Biosystems*) manufacturer instructions. Electrophoresis was performed in an *ABI3100* automated sequencer (*ABI Prism*® *3100 Applied Biosystems*) and electropherograms were analysed by *SeqScape v2.5 Software* (*Applied Biosystems*). *GENES version 2009.7.0* (Cruz, 1998) and *Arlequin* (Excoffier & Lischer, 2010) was used to obtain allelic and genotypic frequency and Hardy-Weinberg Equilibrium (HWE) estimates, and to perform chi-square tests of frequency comparisons. To reduce the bias from unbalanced sample size among locations, samples were grouped by State (Ceará vs São Paulo) and performance test.

3 RESULTS

Table 1 shows *FecG^E* allelic and genotypic frequencies for individual farms. Overall observed *FecG^E* frequency was 0.65. HWE deviations were observed in samples from Ceará and São Paulo, with an increase number of observed heterozygotes ($\chi^2=7.274$, $P< 0.01$). The same was true for an AMOVA test were 7% ($p<0.001$) of genetic difference was observed between both States. *FecG^E* genotypic and allelic frequencies for performance-tested rams classified according to performance test classification are shown in Table 2. Observed *FecG^E* allelic frequencies in samples from Ceará, São Paulo and overall were 0.76, 0.70 and 0.75, respectively and they do not departed from HWE. In addition, significant differences in *FecG^E* allelic frequencies between Elite/Superior (0.76) and Regular/Inferior (0.70) groups were not observed ($\chi^2= 1.98$, $P=0.3708$). Logistic regression with Ceara PT samples showed a moderate association of homozygous *FecG^E* (geno=1- Figure 1) for animals classified as Regular/Inferior (class 0- Figure 1) with an odds ratio of 0.463.

Table 1. *FecG^E* allele and genotype frequencies by farm for Morada Nova sheep from Ceará (N= 140) and São Paulo (N= 112) States, Brazil.

		Ceará																					São Paulo						
Farms		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	Total	A	B	C	D	E	Total
Genotypic Frequency*	<i>FecG_E^{E/}</i>	0.3	0.4	0.4	0.4	0.6	0.7	0.6	0.8	0.8	1	0.8	0.6	0.7	0.3	0	1	1	0.3	0.7	0.5	0.3	0.53	0.6	0.1	0	0.3	0.1	0.2
	<i>FecG₊^{E/}</i>	0.7	0.6	0.5	0.6	0.3	0.3	0.3	0.3	0.2	0	0.2	0.2	0.3	0.7	1	0	0	0.7	0.2	0.5	0.7	0.42	0.4	0.7	1	0.7	0.8	0.67
	<i>FecG_{/+}⁺</i>	0	0	0.1	0	0.1	0	0.1	0	0	0	0	0.2	0	0	0	0	0	0	0.1	0	0.1	0.05	0.1	0.2	0	0	0.2	0.13
Number Animals		7	7	17	5	8	9	10	3	13	1	5	5	10	3	1	1	1	3	9	4	18	140	20	32	1	23	36	112
Allele Frequency*	<i>FecG⁺</i>	0.4	0.3	0.4	0.3	0.3	0.2	0.3	0.2	0.1	0	0.1	0.3	0.2	0.3	0.5	0	0	0.3	0.2	0.3	0.4	0.26	0.3	0.5	0.5	0.4	0.6	0.46
	<i>FecG^E</i>	0.7	0.7	0.6	0.7	0.8	0.8	0.8	0.8	0.9	1	0.9	0.7	0.9	0.7	0.5	1	1	0.7	0.8	0.8	0.6	0.74	0.8	0.5	0.5	0.6	0.4	0.54

*Difference between states was significant $p < 0.01$.

Table 2. *FecG^E* genotypic and allelic frequencies in Morada Nova rams in Ceará (N= 122) and São Paulo (N= 17) States according to performance test classification.

Performance test	Genotypes	Frequency (N)	Alleles	Frequency	Elite/ Superior	Regular/ Inferior
2008 (CE)	+/ +	0.03 (1)	<i>FecG⁺</i>	0.20	0.10 (1)	0.00 (0)
	+/E	0.33 (9)	<i>FecG^E</i>	0.80	0.10 (1)	0.47 (8)
	E/E	0.64 (18)	-		0.80 (9)	0.53 (9)
2009-1 (CE)	+/ +	0.08 (2)	<i>FecG⁺</i>	0.32	0.17 (2)	0.00 (0)
	+/E	0.48 (12)	<i>FecG^E</i>	0.68	0.33 (4)	0.62 (8)
	E/E	0.44 (11)	-		0.50 (6)	0.38 (5)
2009 -2 (CE)	+/ +	0.03 (1)	<i>FecG⁺</i>	0.23	0.00 (0)	0.05 (1)
	+/E	0.41(14)	<i>FecG^E</i>	0.77	0.37 (6)	0.45 (8)
	E/E	0.56 (19)	-		0.63 (10)	0.50 (9)
2010 (CE)	+/ +	0.06 (2)	<i>FecG⁺</i>	0.23	0.06 (1)	0.05 (1)
	+/E	0.34 (12)	<i>FecG^E</i>	0.77	0.44 (7)	0.27 (5)
	E/E	0.60 (21)	-		0.50 (8)	0.68 (13)
CE (Total)*	+/ +	0.05 (6)	<i>FecG⁺</i>	0.24	0.07 (4)	0.03 (2)
	+/E	0.39 (47)	<i>FecG^E</i>	0.76	0.33 (18)	0.43 (29)
	E/E	0.56 (69)	-		0.60 (33)	0.54 (36)
SP*	+/ +	0.00 (0)	<i>FecG⁺</i>	0.30	0.00 (0)	0.00 (0)
	+/E	0.59 (10)	<i>FecG^E</i>	0.70	0.55 (5)	0.63 (5)
	E/E	0.41 (7)	-		0.45 (4)	0.37 (3)
Total**	+/ +	0.04 (6)	<i>FecG⁺</i>	0.25	0.06 (4)	0.03 (2)
	+/E	0.41 (57)	<i>FecG^E</i>	0.75	0.36 (23)	0.44 (34)
	E/E	0.55 (76)	-		0.58 (37)	0.52 (39)

*Difference between States was non-significant.

** Difference between Performance Test Classification was non-significant.

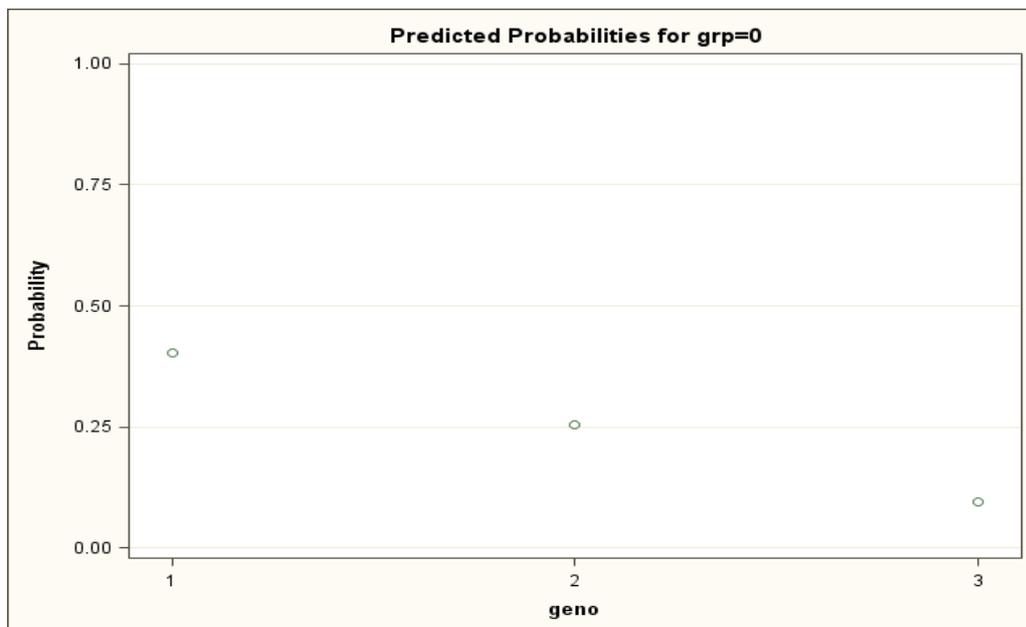


Figure 1- Logistic regression of genotypes *FecG^E/FecG^E* (1), *FecG^E/FecG⁺* (2) and *FecG⁺/FecG⁺* (3) for Morada Nova animals classified as Regular/inferior in performance test (grp=0).

4 DISCUSSION

Previous *GDF9* studies in sheep (Chu et al., 2011b; Mullen & Hanrahan, 2014; Souza et al., 2014) and goat breeds (Ahlawat et al., 2013) with a high frequency of multiple ovulations, did not observe the *FecG^E* allele after its discovery by Silva et al. (2011). In addition, analyses of whole genome shotgun NGS data generated by the International Sheep Genome Consortium from 75 animals from 39 different breeds and two wild species (*Ovis canadensis* and *Ovis dalli*) did not reveal any *FecG^E* alleles in any of the sequenced animals (data not shown), corroborating claims from Silva et al. (2011) that this mutation was first identified in Brazilian hair sheep. Nevertheless, additional ovulation studies are required to corroborate the present hypothesis.

The results represent the first *FecG^E* frequency estimates reported for Morada Nova sheep and show that this *GDF9* mutation is not restricted to Santa Inês hair sheep, as previously reported (Silva et al., 2011). In fact, observed *FecG^E* frequencies in Morada Nova (0.74) were significantly higher than reported frequencies for Santa Inês (0.52, Silva et al., 2011). Considering the effect of *FecG^E*, these results are corroborated by average litter size comparisons between the two breeds. While mean litter sizes of 1.25 have been reported in Santa Inês (Mexia et al., 2004), litter sizes ranging from 1.37 to 1.48 have been observed in Morada Nova (Machado, 1999; Quesada, et al., 2002; Selaive-Villarroel & Fernandes, 2000). Analyses of the genetic relationship between the two breeds suggest that Santa Inês was probably derived by alternate crossbreeding between Morada Nova, Brazilian Bergamacia and other locally adapted sheep breeds from Brazil (McManus et al., 2010).

The high number of Morada Nova animal homozygote and heterozygote for *FecG^E* observed in farms in Ceará and São Paulo States, and previous studies with Santa Inês, suggest that this allele may be found in most Brazilian locally adapted sheep breeds. In addition, low levels of introgression of *FecG^E* have also been observed in commercial or specialized

breeds in Brazil (data not shown). These observations reveal a new possible approach for conservation and breeding of Brazilian sheep breeds as genotyping and selection for this *GDF9* variant may be done at low costs with great potential impacts to local production systems. Farmers with extensive production systems in extreme environments may select against *FecG^E* while those with semi-intensive systems and better pastures can select in favor of this allele to increase twin birth rates in their flocks.

The significant difference of *FecG^E* allelic frequencies observed between samples from Ceará and São Paulo States suggests distinct selection practices may be in place in different regions. Litter size has been shown to significantly affect farm profitability, but this may be highly influenced by production system and region. Energy requirements from ewes with multiple lambs may not be fully met in low input systems frequently observed in Brazil's semi-arid region, therefore resulting in higher lamb mortality rates and lower profits as a consequence. Selection for multiple lambing may therefore lead to higher profits only in production systems using higher inputs usually seen in Southeastern Brazil. Conversely, founder effects and drift may represent significant factors to explain observed *FecG^E* allelic frequencies differences as studied flocks have small effective numbers.

Litter size was identified as one of the most important traits for determining overall profitability of lamb production with Morada Nova sheep in Brazilian production systems (Lôbo et al., 2011), but current performance tests do not consider this trait for computing final scores and ranking. Rao & Notter (2000) did not find any strong genetic correlations between litter size and different growth traits. Therefore, current performance tests based solely on growth and carcass traits are not likely to directly or indirectly contribute to increase prolificacy in Morada Nova sheep. The overall observed frequency of *FecG^E* was 0.75 and differences in homozygote frequencies in the Elite/Superior and Regular/Inferior groups were not significant (Table 2, Figure 1). These results indicate that *FecG^E* allele is widespread in Morada Nova sheep and that inclusion of genotyping results in final performance test scores would contribute to increase allele frequencies and therefore, litter size and overall profitability of lamb production with this breed. Future studies to derive distinct indexes, fine-tuned to the production system used (intensive and extensive), may greatly improve genetic gains obtained in line with selection objectives based on regional farm profitability. Moreover, this study is important to connect the community-based breeding program with the National conservation program. There is urgency in adding germplasm with both *FecG^E* and wild type alleles in gene banks to assure food security and phenotypic plasticity to the activities performed by private sector.

5 CONCLUSION

Sheep on the community-based breeding program for Morada Nova hair sheep showed a high allelic diversity for the *FecG^E* allele (*GDF9* gene), which has been shown to increase litter size. Genetic characterization for this allele in young rams on performance tests may contribute to increase litter size in the breed, a trait shown to affect farm profitability. Genetic monitoring for this SNP mutation in the *GDF9* gene may be useful in the breeding program even if prolificacy is an inherently female characteristic. This allele had previously been described only in Brazilian Santa Ines sheep and results from this study suggest that this mutation might be associated with Brazilian locally adapted sheep breeds and can be used to improve sheep production in the country.

6 REFERENCES

- AHLAWAT, S.; SHARMA, R.; MAITRA, A. Screening of indigenous goats for prolificacy associated DNA markers of sheep. **Gene**, v. 517, n. 1, p. 128–131, 2013.
- CASTANHEIRA, M. et al. Use of heat tolerance traits in discriminating between groups of sheep in central Brazil. **Tropical Animal Health and Production**, v. 42, n. 8, p. 1821–1828, 2010.
- CHU, M. X. et al. GDF9 as a candidate gene for prolificacy of Small Tail Han sheep. **Molecular biology reports**, v. 38, n. 8, p. 5199–204, 2011.
- CRUZ, C. D. **Programa GENES: Aplicativo Computacional em Estatística Aplicada à Genética (GENES - Software for Experimental Statistics in Genetics) Genetics and Molecular Biology** scielo , , 1998.
- DOMINGUES, O. **Sobre a origem do carneiro deslanado do Nordeste**. Fortaleza Publicação numero: 3, , 1954.
- EXCOFFIER, L.; LISCHER, H. E. L. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. **Molecular Ecology Resources**, v. 10, n. 3, p. 564–567, 2010.
- FACÓ, O. et al. Raça Morada Nova: Origem , Características e. In: Sobral: Documentos / Embrapa Caprinos, ISSN 1676-7659 ; 75, 2008. p. 43.
- FACÓ, O. et al. **Teste de Desempenho Individual de Reprodutores da Raça Morada Nova: Resultados da Prova em Morada Nova - CE - 18/02 a 04/06/2008**. Sobral: [s.n.]. Disponível em: <<http://www.cnpc.embrapa.br>>.
- FERNANDES, A. A. O. **Genetic and phenotypic parameter estimates for growth, survival and reproductive traits in Morada Nova hair sheep**. [s.l.] Oklahoma State University, 1992.
- HANRAHAN, J. P. et al. Mutations in the genes for oocyte-derived growth factors GDF9 and BMP15 are associated with both increased ovulation rate and sterility in Cambridge and Belclare sheep (*Ovis aries*). **Biol Reprod**, v. 70, n. 4, p. 900–909, 2004.
- JUENGEL, J. L.; DAVIS, G. H.; MCNATTY, K. P. Using sheep lines with mutations in single genes to better understand ovarian function. **Reproduction**, v. 146, n. 4, p. R111–23, 2013.
- LÔBO, R. N. B. et al. Economic values for production traits of Morada Nova meat sheep in a pasture based production system in semi-arid Brazil. **Small Ruminant Research**, v. 96, n. 2-3, p. 93–100, 2011.
- MACHADO, T. M. M. Comportamento Reprodutivo de Ovinos Deslanados no Município de Uberlândia -MG. **Rev. Bras. Reprod. Anim.**, v. 23, p. 189–191., 1999.
- MCMANUS, C. et al. Genetic factors of sheep affecting gastrointestinal parasite infections in the Distrito Federal, Brazil. **Veterinary Parasitology**, v. 166, p. 308–313, 2009.
- MCMANUS, C. et al. Geographical distribution of sheep breeds in Brazil and their relationship

- with climatic and environmental factors as risk classification for conservation. **Brazilian Journal of Science and Technology**, v. 1, p. 3, 2014.
- MCMANUS, C.; PAIVA, S. R.; ARAÚJO, R. O. DE. Genetics and breeding of sheep in Brazil. **Revista Brasileira de Zootecnia**, v. 39, p. 236–246, 2010.
- MEXIA, A. A. et al. Desempenhos reprodutivo e produtivo de ovelhas Santa Inês suplementadas em diferentes fases da gestação. **Revista Brasileira de Zootecnia**, v. 33, n. 3, p. 658–667, 2004.
- MILLER, S. A.; DYKES, D. D.; POLESKY, H. F. A simple salting out procedure for extracting DNA from human nucleated cells. **Nucleic Acids Research**, v. 16, n. 3, p. 1215, 1988.
- MULLEN, M. P.; HANRAHAN, J. P. Direct evidence on the contribution of a missense mutation in GDF9 to variation in ovulation rate of Finnsheep. **PloS one**, v. 9, n. 4, p. e95251, 2014.
- NICOL, L. et al. Homozygosity for a single base-pair mutation in the oocyte-specific GDF9 gene results in sterility in Thoka sheep. **Reproduction**, v. 138, n. 6, p. 921–933, 2009.
- OTSUKA, F.; MCTAVISH, K. J.; SHIMASAKI, S. Integral role of GDF-9 and BMP-15 in ovarian function. **Molecular Reproduction and Development**, v. 78, n. 1, p. 9–21, 2011.
- PAIM, T. P. et al. Estudo econômico da produção de cordeiros cruzados confinados abatidos em diferentes pesos. **Ciência Animal Brasileira**, v. 12, p. 48–57, 2011.
- QUESADA, M.; MCMANUS, C.; COUTO, F. A. D. Estimation of genetic and phenotypic parameters for Brazilian hair sheep in Central Brazil. **Efeitos genéticos e fenotípicos sobre características de produção e reprodução de ovinos deslanados no Distrito Federal**, v. 31, n. 1 SUPPL., p. 342–349, 2002.
- RAO, S.; NOTTER, D. R. Genetic analysis of litter size in Targhee, Suffolk, and Polypay sheep. **Journal of Animal Science**, v. 78, n. 8, p. 2113–2120, 2000.
- SELAIVE-VILLARROEL, A.; FERNANDES, A. Desempenho reprodutivo de ovelhas deslanadas Morada Nova no estado do Ceará. **Revista Científica de ...**, 2000.
- SILVA, B. D. M. et al. A new polymorphism in the Growth and Differentiation Factor 9 (GDF9) gene is associated with increased ovulation rate and prolificacy in homozygous sheep. **Animal Genetics**, v. 42, n. 1, p. 89–92, 2011.
- SOUZA, C. J. H. et al. Mutation in the protease cleavage site of GDF9 increases ovulation rate and litter size in heterozygous ewes and causes infertility in homozygous ewes. **Animal genetics**, v. 45, n. 5, p. 732–9, 2014.

CAPÍTULO 3

MINERAÇÃO DE POLIMORFISMOS DE BASE ÚNICA (SNPs) EM GENES RELACIONADOS À TAXA DE OVULAÇÃO EM OVINOS

RESUMO

Como resultado de estudos em diferentes raças de ovinos no Brasil e no mundo, mutações foram descobertas em genes associados ao controle da foliculogênese, o que no final pode afetar na fertilidade e taxa de ovulação. A identificação de sítios polimórficos em genes associados à taxa de ovulação realizados nesse trabalho irão compor painéis ligados à prolificidade em ovinos. Sequências completas de três genes relacionados à taxa de ovulação: *BMP15*, *BMPR1-b* e *GDF9* foram obtidas por dados de sequenciamento de nova geração do genoma inteiro de 75 ovinos de diferentes raças provenientes do repositório do Consórcio Internacional do Genoma Ovino (ISGC). Este estudo mostra a identificação de 1610 SNPs em genes relacionados à prolificidade em ovinos, baseado em sequenciamento de genes completo (Pop 1) e de fragmentos do exon (12 SNPs) (Pop 2). Dos SNPs identificados na Pop 1 (ISGC, N=75), 36 se encontram em exons dos genes *GDF9* (13), *BMP15* (7) e *BMPR1-b* (16). Onze destes SNPs (6 no gene *GDF9*, 2 no gene *BMPR1-b*, e 3 no gene *BMP15*) já foram descritos anteriormente na literatura em diferentes raças deslanadas e lanadas de ovinos. O SNP 488Br encontrado no gene *BMPR1-b*, associado a fenótipos de hiperprolificidade em raças de ovinos, foi observado pela primeira vez nas raças prolíficas: Bangladeshi, Garut e Sumatra. Dois SNPs que provocam alterações não conservativas nos AAs de *BMP15*, foram pela primeira vez relatados em quatro animais de raças consideradas prolíficas; Cheviot, Churra e Swiss White Alpine. Desvio de EHW ($0,05 > P < 0,001$) foram identificados em 634 SNPs, enquanto 1404 SNPs apresentaram valor de $MAF > 0,01$ nas sequências dos três genes da população 1. Foram estimados haplótipos utilizando SNPs com mutações não conservativas, e observamos haplótipos com presença de duas ou mais variantes alélicas entre genes diferentes. A identificação e seleção dos SNPs encontrados no atual trabalho, serão no futuro utilizados para compor um painel de baixa densidade para a prolificidade e após análises de validação em raças prolíficas, com histórico de parto simples e múltiplos, poderão ser usados para otimizar genotipagem dos animais pertencentes aos programas de conservação e melhoramento de ovinos.

Palavras-chave: prolificidade, GDF-9, BMP15, BMPR1-b, *ovis airies*, sequenciamento de nova geração, seleção assistida por marcadores

ABSTRACT

SCREENING THE SINGLE NUCLEOTIDE POLYMORPHISMS (SNPS) IN GENES RELATED TO OVULATION RATE IN SHEEP

Genetic studies using different breeds of Brazil and around the world have found mutations in genes associated with the control of folliculogenesis, which is related to fertility and ovulation rate. The identification of polymorphic sites in genes associated with fertility aims to compose small panels associated to prolificacy in sheep. We isolated the complete sequence of three genes related to ovulation rate generated by Next Generation Sequence (NGS); *BMP15*, *BMPRI-b* and *GDF9* of 75 individuals from the International Sheep Genome Consortium (ISGC) repository. The search for polymorphisms using bioinformatics tools in the sequences resulted in 48, 101 and 1461 polymorphic sites, 19, 69 and 1314 and located in intron 13, 6, and 16 in the exons of the genes *GDF9*, *BMP15* and *BMPRI-b*, respectively. Eleven of these SNPs (6 in *GDF9* gene, 2 in *BMPRI-b* gene, and 3 in *BMP15* gene) were previously described in the literature in different hair breeds sheep. Noteworthy is the SNP 488Br in the *BMPRI-b* gene, associated with hyper prolific breeds, this SNP was observed for the first time in the prolific breeds; Bangladeshi, Garut, and Sumatra. Two SNPs that cause non-conservative changes in AAs of the gene *BMP15* were first time reported in four animal breeds considered prolific; Cheviot, Churra and Swiss Alpine White. HWE disequilibrium ($0.05 > P < 0.001$) were identified in 634 SNPs, and 1404 SNPs with $MAF > 0.01$ were found in the sequences of the three genes. Estimates of haplotypes were calculated using non-conservative mutations, and was observed the presence of haplotypes with with two or more allelic variants in different genes. The selection of SNPs identified in the current work will be useful in the future to compose a low-density panel for prolificacy and after that; the panel should be validated in prolific breeds with history of simple and multiple parturition. As a final goal, the panel should be used for routinely genotyping of animals belonging to conservation programs and breeding programs.

Keywords: Prolificacy, GDF-9, BMP15, BMPRI1b, *ovis aries*, next generation sequence, markers assisted selection (MAS).

1 INTRODUÇÃO

Muitos trabalhos de prospecção de polimorfismos de base única (SNPs) em genes que afetam características de interesse econômico vem sendo executados desde a década de 90 por diferentes grupos de pesquisa no mundo (ex., Rothschild & Soller, 1997; Meira et al., 2014; Goddard & Hayes 2009). Estes estudos foram, em geral, baseados em marcadores moleculares associados a um QTL, marca de seleção de interesse para a detecção de genes, e mutações causais responsáveis pelos efeitos observados (Borevitz & Chory, 2004; Das et al., 2014; Ernst & Steibel, 2013; Purvis & Franklin, 2005; Sahana et al., 2014). É importante reforçar que a maioria das características de interesse pecuário é controlada por muitos genes de efeito pequeno, o que dificulta a busca das mutações causais responsáveis pela maior variação do fenótipo de interesse (Andersson & Georges, 2004; Goddard & Hayes, 2009). Além disso, as tecnologias disponíveis há 10-15 anos atrás eram restritas à um número pequeno de marcadores e a populações experimentais pequenas, compostas de centenas de animais tipicamente estruturados em famílias segregantes.

Apenas como exemplo, as técnicas padronizadas para prospecção de SNPs eram baseadas na metodologia de sequenciamento de Sanger (Calvo et al., 2006; De Camargo et al., 2012; Fahrenkrug et al., 2002; Veneroni et al., 2010), na qual o custo por pares de bases, bem como no tempo necessário para obtenção dos resultados é infinitamente maior (Ansorge, 2009).

O desenvolvimento das tecnologias de sequenciamento de nova geração (NGS) representadas principalmente pelas plataformas: 454 FLX da Roche (Margulies et al., 2005), Solexa da Illumina (Bennett, 2004) e a SOLiD System da Applied Biosystems (Valouev et al., 2008), permitiu a disponibilização de um grande conjunto de dados a partir de um único experimento, reduzindo custos e acelerando a velocidade da pesquisa. Esses dados vêm permitindo a identificação de vários polimorfismos de sequências que podem ser incorporados

em estudos visando uma melhor compreensão das características poligênicas, consequentemente, a aplicação de novas estratégias de seleção em programas de melhoramento. O desenvolvimento de novas tecnologias para obtenção de sequências de DNA permitiu, paralelamente, o desenvolvimento acelerado de metodologias de genotipagem em larga escala a um baixo custo dos SNPs encontrados (Kijas et al., 2012).

Estudos vêm sendo realizados com sucesso para busca e prospecção de SNPs em genes de interesse econômico usando a tecnologia sequenciamento de nova geração em plantas (Barchi et al., 2011; Kumar et al., 2012; Zou et al., 2014) e animais (Barris et al., 2012; Fontanesi et al., 2015; Heaton et al., 2014). Heaton et al. (2013), a partir de sequências dos genomas completos de 75 indivíduos do Consórcio Internacional de Genoma de Ovinos (ISGC – www.sheepmap.com), por exemplo, identificaram polimorfismos de sequências em regiões codantes de um gene associado à suscetibilidade/resistência da infecção do vírus da pneumonia progressiva ovina (OPPV), que foram utilizados para compor painéis relacionados à resistência dessa infecção (OPPV) em ovinos. De maneira geral, podemos dizer que com a finalização do sequenciamento do genoma completo de ovinos (Jiang et al., 2014) o processo de identificação de SNPs para a espécie foi simplificado.

A abordagem do presente capítulo é similar à de Heaton et al. (2013), onde foram usados os mesmos dados de sequências geradas pelo ISGC (www.sheepmap.com) de 75 indivíduos de diferentes raças, sendo que, no presente estudo, o objetivo foi identificar SNPs nos principais genes associados à taxa de ovulação: *Bone morphogenetic proteins-15 (BMP15)*, *Bone morphogenetic protein receptor type 1b (BMPRI-b)* e *Growth and differentiation factor (GDF9)* e suas regiões flanqueantes. Essa procura, visa em última instância, auxiliar o desenvolvimento de painéis de marcadores SNP de baixa densidade para testes diagnósticos simplificados.

2 MATERIAIS E MÉTODOS

2.1 Material biológico e amostragem

2.1.1 Amostras do consórcio Internacional do genoma Ovino (População1).

Foram utilizadas sequências geradas via NGS do genoma completo de 75 ovinos provenientes do repositório do Consórcio Internacional do Genoma Ovino – (ISGC – www.sheephapmap.com) (Tabela 1). As sequências dos genes *BMP15*, *BMPRI-b* e *GDF9* foram obtidas por alinhamento de curtos segmentos de leitura (*reads*), com média de ≈ 100 pb de comprimento contra o genoma de referência Oar3.0 (<http://www.livestockgenomics.csiro.au/sheep/oar3.0.php>) (montagem disponível no ano 2012, quando foram geradas as sequências). Cada um dos 75 animais gerou em média um total de 38,4 Gb com coberturas de 10X.

Os animais pertencentes ao ISGC são de 39 raças e duas espécies selvagens (*Ovis canadensis* e *Ovis dalli*) dos quais seis eram representantes de três raças brasileiras: dois ovinos da raça Crioula Lanada, dois da raça Morada Nova, dois da raça Santa Inês (Tabela 1). A escolha dos animais pelo Consórcio, baseou-se no trabalho de Kijas et al. (2012) em que foi avaliada a distribuição da variabilidade genética e o impacto da seleção após a domesticação em 2.819 animais de 74 raças ovinos a partir da genotipagem de mais de 49.000 SNPs contidos no *Illumina ovineSNP50Beadchip* (Illumina Inc., San Diego, Califórnia, EUA).

2.1.2 Amostras de ovinos das raças brasileiras: Morada Nova e Santa Inês (População2).

A População2 (N=436) é composta de duas raças localmente adaptadas brasileiras: Morada Nova (OMN) e Santa Inês (OSI). A raça Morada Nova com N = 250 contém 111 animais do programa de melhoramento da Embrapa Caprinos e Ovinos, e dentre estes, 23 ovinos possuem informação de genotipagem *OvineSNP50 Genotyping BeadChip - Illumina* (Illumina Inc., San Diego, CA) usados no trabalho de Kijas et al. (2012), e sendo dois (BMN3 e BMN4) selecionados para compor a população dos 75 ovinos do Consórcio Internacional. Os 139 animais restantes eram carneiros jovens que participaram de quatro testes de desempenho (entre 2008 e 2010 - capítulo 2), distribuídos em fazendas de dois estados brasileiros, Ceará (122 animais) e São Paulo (17 animais) (Facó et al., 2009).

A raça Santa Inês (OSI) com N= 186 possui 49 animais que também foram usados para genotipagem de 50K *Illumina OvineSNP50 beadchip* (Illumina Inc, San Diego, CA) (Kijas et al., 2012), dois ovinos dos quais (BSI3 e BSI4) foram selecionados para o experimento de sequenciamento do genoma completo. Adicionalmente, 137 ovinos da raça Santa Inês que participam do programa de conservação da Embrapa Tabuleiros Costeiros (CPATC) foram utilizados.

As amostras de DNA desses animais se encontravam estocadas no banco de DNA do laboratório de Genética Animal (LGA) da Embrapa Recursos Genéticos e Biotecnologia (CENARGEN). A população 2 foi inserida nesse estudo para servir como amostra adicional para verificar a frequência de possíveis SNPs identificados no exon 2 do gene *GDF9*. Esse exon/gene foi verificado mais profundamente em razão de possuir uma mutação de interesse para ovinos deslanados identificada por Silva et al. (2011).

Tabela 1- Os 75 animais de 39 raças e duas espécies selvagens que serão usados para a análise de sequências dos genes associados à alta taxa de ovulação, provenientes do repositório do Consórcio Internacional do Genoma Ovino (ISGC).

Raças	País de origem*	Aptidão*	Prolificidade (TN)**	Tipo*	Número de animais por raça (N)
African White Dorper	África do Sul	Carne	Prolífica	Deslanada	2
Afrikaner, Namaqua	África do Sul	Carne, couro	Prolífica	Deslanada	1
Afrikaner, Ronderib	África do Sul	Carne, couro	Não informado	Deslanada	2
Afshari	Irã	Carne	Prolífica	Lanada	2
Awassi	Oriente medio	Leite	Não prolífica	Lanada	1
Awassi, Turkish	Turquia	Leite, carne e lã	Não prolífica	Lanada	2
Bangladeshi	Bangladesh	Lã	Prolífica	Lanada	2
Brazilian Creole	Brasil	Lã	Prolífica	Lanada	2
Castellana	Espanha	Carne, leite, lã	Prolífica	Lanada	2
	Jammu e Caxemira		Não prolífica		
Changthangi	(Índia)	Couro, carne, lã		Lanada	2
Cheviot	Inglaterra-Escócia	Carne, lã	Prolífica	Lanada	2
Churra	Espanha	Carne, leite	Prolífica	Lanada	2
Cine Capari	Turquia	Carne, Leite	Não prolífica	Lanada	1
Dorset, Poll	EUA	Carne	Prolífica	Lanada	1
Ethiopian Menz	Etiópia	Carne	Não prolífica	Lanada	1
Finnsheep	Finlândia	Carne	Prolífica	Lanada	2
Garole	Bangladesh	Carne	Prolífica	Lanada	1
Garole, Indian	Índia	Carne	Prolífica	Lanada	1
Garut	Indonésia	Carne	Prolífica	Lanada	2
Gulf Coast native	Estados Unidos	Carne, lã	Prolífica	Lanada	2
Karakas	Turquia	Lã	Não prolífica	Lanada	2
Karya	Turquia	Carne, leite	Prolífica	Lanada	1
Lacaune, Meat	França	Carne	Prolífica	Lanada	1
Lacaune, Milk	França	Carne, leite	Prolífica	Lanada	1
Merino	Espanha	Carne, lã	Prolífica	Lanada	3
Morada Nova	Brasil	Carne, Couro.	Prolífica	Deslanada	2
Norduz	Turquia	Lã	Não prolífica	Lanada	2
Ojalada	Espanha	Carne	Não prolífica	Lanada	2
<i>Ovis canadensis</i>	América do Norte	Espécie selvagem	Não prolífica	Deslanada	3
<i>Ovis dalli</i>	América do Norte	Espécie selvagem	Não prolífica	Deslanada	2
Romney	Reino Unido	Carne, lã	Prolífica	Lanada	1
Sakiz	Turquia	Lã	Prolífica	Lanada	2
Salz	Espanha	Carne	Prolífica	Lanada	3
Santa Inês	Brasil	Carne	Prolífica	Deslanada	2
Scottish Blackface	Escócia	Carne, lã	Prolífica	Lanada	1
Sumatra	Sumatra	Carne	Prolífica	Lanada	2
Swiss Mirror	Suíça	Carne	Prolífica	Lanada	1
Swiss White Alpine	Suíça	Carne, lã	Prolífica	Lanada	4
Texel	Holanda	Carne, lã	Prolífica	Lanada	1
Tibetan, Eastern	Tibet	Carne	Não prolífica	Lanada	1
Tibetan, Northern	Tibet	Carne	Não prolífica	Lanada	1
Valais Blacknose	Suíça	Carne	Prolífica	Lanada	1
Welsh Hardy Speckled Face	País de Gales	Carne	Não informado	Lanada	1
Welsh Mountain, Dollgellau	País de Gales	Carne	Não prolífica	Lanada	1
Welsh mountain, Tregaron	País de Gales	Carne	Não prolífica	Lanada	1

*Informações sobre aptidão, prolificidade, país de origem e tipo de raça foram adquiridas pelo banco de dados *on line* <http://dad.fao.org>.

**Informação de prolificidade foi avaliado pela média do tamanho da ninhada (TN), considerando prolíficos aqueles com a TN \geq 1,

2.2 Montagem das sequências completa dos genes candidatos selecionados utilizando a População 1

Os dados recebidos do Consórcio Internacional do Genoma Ovino (ISGC) foram previamente alinhados com genoma de referência ovino versão 3.0 (Oar3.0) <http://www.livestockgenomics.csiro.au/sheep/oar3.0.php>, usando a ferramenta *Burrows Wheeler Alignment* (BWA) (Li et al., 2009) para gerar uma sequência referência própria da população dos 75 indivíduos do ISGC na região dos três genes reconhecidamente relacionados a taxa de ovulação (*GDF9*, *BMP15* e *BMPRI-b*).

Em razão da presença de muitas lacunas (*gaps*) na versão do genoma ovino usada, foi empregada uma estratégia adicional de fazer uma montagem *de novo* da sequência de referência dos três genes em estudo (*GDF9*, *BMP15* e *BMPRI-b*). Deste modo, as sequências dos três genes de interesse de cada animal foram extraídas utilizando *SAMtools* (Sequence alignment/map) (Li et al., 2009) e convertidas para o formato *fastQ* usando o *bam2fastQ*, ferramenta do repositório *BamUtil* (<http://genome.sph.umich.edu/wiki/BamUtil>). Em seguida foi feita uma montagem *de novo* comparando os 75 animais do Consórcio usando o programa *velveth* (Zerbino & Birney, 2008).

A remontagem gerou uma nova sequência referência, denominada REFISGC, de cada um dos três genes: *BMP15*, *GDF9* e *BMPRI-b*, abrangendo 2000 pb antes e depois da região codificadora. Quando as sequências REFISGC de cada gene foram alinhadas contra o genoma referência, a versão referência atualizada utilizada foi a Oar3.1 (<http://www.livestockgenomics.csiro.au/sheep/oar3.1.php>). Posteriormente foi realizado um alinhamento contra as sequências depositadas no *RefSeq sequence Genebank*: NC_019462.1 (*GDF9*), NC_019463.1 (*BMPRI-B*) e NC_019484.1 (*BMP15*), via *BLASTN* v.2.2.312 (*Basic Local Alignment Search Tool*) (Zhang et al., 2000). Os limites entre introns e exons em cada gene sequenciado foi obtido novamente via *BLASTN* v.2.2.312 contra as sequências de mRNA depositadas no também no *GeneBank*; NM_001142888.2 (*GDF9*), NM_001114767.1 (*BMP15*) e NM_001009431.1 (*BMPRI-B*).

Estas sequências foram consideradas como a referência para o alinhamento das sequências gênicas de todos 75 indivíduos que, por sua vez, foram alinhadas usando o software BWA (Li et al., 2009) a partir da nova sequência referência REFISGC de cada gene. Os arquivos, depois de alinhados, foram transformados no formato BAM e a indexação do arquivo foi realizada usando o programa *SAMtools* (Sequence alignment/map) (Li et al., 2009). A

visualização da montagem das sequências de cada animal foi realizada a partir do programa IGV (Integrative Genomics Viewer) (Robinson et al., 2011).

Os módulos *SeqManGen* e *Megalign* do software *DNASTAR Navigator* (DNASTAR Inc., Madison, WI. USA) foram usados para conversão dos diferentes formatos de sequências (*.bam; *.bami e *.fasta) bem como para realizar os alinhamentos; e o software *MEGA* (Tamura et al., 2013) foi usado para realinhamentos das sequências gênicas em fasta dos 75 indivíduos.

2.3 Identificação de SNPs na População1 para os três genes selecionados.

A identificação dos SNPs nas sequências dos três genes *GDF9*, *BMP15* e *BMPR1b*, geradas por NGS, foi realizada pelo programa *SamTools* com o comando *mpileup* (Li et al., 2009), que posiciona os alinhamentos de todos os segmentos de leitura a partir da sequência referência. Este programa também estima todos os possíveis genótipos gerados pelos segmentos de leitura e calcula a probabilidade do verdadeiro genótipo de cada amostra, baseado em algoritmo que calcula a probabilidade de um provável polimorfismo vir a ser um erro de sequenciamento. Os resultados foram gerados no formato BCF (*binary call format*), que por sua vez foram usados para identificação dos SNPs e InDels a partir do comando *bcftools*. As variantes identificadas foram exportadas em um arquivo de formato VCF versão 4.0 (*variant call format*), gerados a partir de scripts customizados em PERL.

Os arquivos VCF continham as seguintes informações de cada SNP: identificação (ID); cromossomo (CHROM); posição (POS); alelo da sequência da referência (REF); o provável polimorfismo (ALT); e qualidade do SNP calculado pela probabilidade de todas as amostras serem homozigotos conforme a referência baseados na escala Phred (QUAL). Adicionalmente, o arquivo gerado exportou uma lista de informações (INFO) sobre os SNPs para cada animal contendo: o genótipo (GT); número de segmentos de leitura de alta qualidade baseado na cobertura de cada SNP detectado (DP) e qualidade genotípica (GQ). A partir dessas informações os SNPs foram filtrados utilizando dois parâmetros: (1) Qualidade do SNP calculado (baseados na probabilidade de escala Phred), (2) índice de cobertura dos segmentos de leitura (DP).

2.4 Amplificação, sequenciamento e alinhamento das sequências gênicas do exon 2 do gene *GDF9*. (População2).

A população2 foi sequenciada para o segundo exon no gene *GDF9* (GB:AF078545) de tamanho de 965 pb (Bodensteiner et al., 1999). Os amplicons foram gerados por dois pares de iniciadores (vide capítulo 2) e o produto obtido foram fragmentos de 400 e 600pb em média. O sequenciamento foi realizado utilizando metodologia Sanger em sequenciador automático *ABI3100* (*ABI Prism*® *3100 Applied Biosystems*) e a rotina foi a mesma descrita no Capítulo 2.

2.5 Análises estatísticas

Parâmetros genéticos de cada SNP incluindo: (1) frequência alélica, (2) frequência de heterozigotos observados (HETObs) e (3) esperados (HETEsp); (4) Equilíbrio de Hardy-Weinberg (EHW); (5) percentual mínimo de genotipagem (%Geno), (6) frequência do alelo menor (MAF) foram obtidos utilizando os programas *Genealex* (Pealkall & Smouse, 2012) e *Arlequin* (Schneider et al., 2000). Essas estimativas foram calculadas na população 1 e 2, separadamente.

Haplótipos dos 75 indivíduos diplóides (150 sequências – Pop1) de cada gene foram estimados utilizando *PHASE v2.1.1 (DNAsp)* (Librado & Rozas, 2009) para todos os genes juntos e, em seguida, para cada gene separadamente.

3 RESULTADOS

3.1 Montagem e alinhamento da sequência completa dos genes *BMP15*, *BMPRI-b* e *GDF9*

As sequências dos genes *BMP15*, *BMPRI-b* e *GDF9* foram montadas por alinhamento de curtos segmentos de leitura (*reads*) obtidas do banco de dados do Consórcio Internacional do Genoma Ovino, com média de ≈ 100 pb de comprimento contra o genoma de referência Oar3.0 (<http://www.livestockgenomics.csiro.au/sheep/oar3.0.php>), resultando nas sequências referências (REFISGC) de cada gene. Foram extraídos fragmentos de sequências mapeados nos cromossomos 5, 6 e X com tamanho de 4544 pb, 233001 pb e 8675 pb, que corresponde aos genes *GDF9*, *BMPRI-b* e *BMP15*, respectivamente (Tabela 2).

A sequência REFISGC do gene *BMPRI-b*, gerada pela montagem dos 75 indivíduos, apresentou diversos *gaps*, já que grande parte das sequências originais da população usada possuíam *gaps*. Portanto, remontagem para uma nova sequência referência foi realizada, utilizando somente os 6 animais brasileiros, selecionados devido à baixa taxa de *gaps* apresentados pelos mesmos. A sequência referência final possui tamanho de 227665 pb e posteriormente foi alinhada com a sequência referência do genoma versão Oar3.1 (<http://www.livestockgenomics.csiro.au/sheep/oar3.1.php>).

A partir da nova sequência referência para o gene *BMPRI-b*, foi possível remontar somente 59 sequências dos 75 animais. As amostras: AWD1, AWD3, AWT1, AW454, BGE4, CHVC1, DWM1, LAC84, MER454, OCAN1, ODAL2, OJA5, SUM7, SALC1, SWA27 e WHSF1 apresentavam regiões com grandes lacunas (*gaps*) impossibilitando a montagem das sequências desses animais de acordo com a nova referência. Já os genes *GDF9* e *BMP15* tiveram 100% das sequências de todos animais montadas.

Os resultados dos alinhamentos dos três genes, gerados por *BLASTN* v.2.2.312 estão apresentados na Tabela 2. A montagem gerada pela sequência do gene *BMPRI-b* foi alinhada a partir da posição 29362087 a 29415836 pb da sequência REFISGC, incluindo parte da sequência do gene *BMPRI-b* contra as sequências do cromossomo 6 (REFOar3.1) depositadas no NCBI. Uma nova sequência REFISGC do gene *BMPRI-b* foi obtida com um tamanho de 54300 pb, eliminando grandes sequências repetidas e considerando a região que abrange o gene e mais 2000pb das regiões flanqueantes.

3.2 Identificação de SNPs nos genes: *GDF9*, *BMPRI-b* e *BMP15* para População1

A mineração por SNPs na Pop1 resultou em um total de 1610 SNPs encontrados nos 3 genes e regiões flanqueantes: *GDF9* (48), *BMP15* (101) e *BMPRI-b* (1461) (Tabela 3). No total, 1402 SNPs estavam presentes em regiões de íntrons, 36 em regiões de exons e ainda 172 em regiões não traduzidas, sendo 116 em 5' UTR e 56 em 3' UTR. Dentre os 36 SNPs localizados na região codificadora, 20 se caracterizaram como mutação conservativas, 15 não conservativas e um códon de parada no gene *GDF9*. A frequência de SNPs encontrados nos genes foram de 1/95pb para o *GDF9*, 1/37pb para o *BMPRI-b* e 1/86 pb para o *BMP15*. Os SNPs foram nomeados por ordem sequencial e com as letras G, B e Br que corresponde aos genes: *GDF9*, *BMP15* e *BMPRI-b*, respectivamente.

No gene *BMPRI-b*, 79% das amostras tiveram suas sequências remontadas, para uma nova busca de SNPs nessas regiões. O processo da busca de SNP nesse gene foi realizado de duas formas: (1) nas 75 sequências resultantes da montagem a partir da sequência referência de regiões com muitos *gaps* e (2) nas 59 sequências resultantes da remontagem a partir da sequência referência com regiões reduzidas de *gap*. Foram encontrados 903 SNPs em (1) e 1347 em (2) e 803 SNPs estavam presentes nas duas análises, ou seja, 120 SNPs únicos em (1) e 538 únicos em (2), totalizando 1461 SNPs, que foram utilizados em análises posteriores.

Tabela 2- Resultados do alinhamento das sequências REFIGSC contra os cromossomos da sequência referência (REFOar3.1[†]) depositada no Genbank (refSeq),

Genes	Id. Sequência	Cobertura (%)	Identidade (%)	Nucleotídeos divergentes (SNPs)	InDel**	Direção do alinhamento	Localização em pb	
							REFISGC	REFoar3.1*
* GDF9	NC_019462.1	100	99	18	5	Reversa	1 a 4544	41840071 a 41844611
** BMP15	NC_019484.1	100	99	10	6	Reversa	1 a 8675	50969838 a 50978518
*** BMPRI-b	NC_019463.1	90	99	41	27	Complementar	1 a 51865	29363564 a 29414140

[†]<http://www.livestockgenomics.csiro.au/sheep/oar3.1.php>

* Sequência referência Oar_v3.1 de *Ovis aries* da raça Texel dos cromossomos :5,6 e X

** Inserção ou deleção de nucleotídeos entre os alinhamentos das respectivas sequências.

O total dos sítios polimórficos nos três genes foi caracterizado em 972 transições, 390 transversões e 248 InDels (inserções/deleções) (Tabela 2). As InDels nas regiões sequenciadas mostraram variantes com regiões repetitivas do mesmo nucleotídeo em 58 indels. Isto pode sugerir erros no *base calling*, um erro comum em dados de sequenciamentos utilizando NGS, assim como InDels heterozigotos em baixa frequência, o que pode refletir erros na montagem das sequências.

A média da qualidade dos SNPs (medida na escala phred) encontrada foi 925,33 (*GDF9*), 925,292 (*BMP15*) e 872,455 (*BMPRI-b*). A média da qualidade dos SNPs no gene *BMPRI-b* foi menor se comparada aos outros genes, o que pode estar associado ao tamanho da sequência, presença de diversas regiões de *gaps*, e regiões duplicadas encontradas na montagem da sequência do gene.

Quinze SNPs localizados em região exônica dos três genes: *GDF9* (7), *BMP15* (3) e *BMPRI-b* (5) e os mesmos já foram previamente identificados e relatados na literatura em diferentes raças, estando eles associados ou não a específicos fenótipos (Tabela 4).

Foram encontrados 51 SNPs monomórficos entre os 75 animais, no entanto polimórficos em relação à sequência REFISGC dos três genes. Dentre os SNPs polimórficos (1559) nos 75 animais 966 (62%) apresentaram MAF acima de 0,05. O percentual de sítios polimórficos entre as raças variou de 3% (N = 50) para raça Garole a 42% (N = 684) para a raça Churra, porém, o efeito de baixa amostragem da maioria das raças (N < 4) pode estar influenciando esse percentual.

Os sítios polimórficos 33G, 146Br, 442Br, 489Br, com alteração conservativa e o InDel 11-12B, alteração não conservativa no aminoácido correspondente, e localizados em regiões de exon, apresentaram MAF > 0,28 (tabela 4) presentes em mais de 21 raças da Pop1 (Tabela 5, 6 e 7).

Todas as raças apresentaram SNPs privativos, ou seja, aqueles encontrados exclusivamente em uma raça, variando de apenas um SNP para a raça Garole (no gene *GDF9* a 25 (nos três genes) para a raça Chura (Tabelas 5, 6 e 7). A espécie selvagem *O. dalli* e mais sete raças: Cheviot, Scottish Blackface, Ethiopian Menz, Karya, Garut, Morada Nova e Changthangi apresentaram SNPs específicos em região de exons de um ou mais dos três genes, sendo cinco mutações conservativas e seis não conservativas (Tabelas 5, 6 e 7).

Tabela 3- Distribuição por classe e localização dos SNPs encontrados após busca de variantes nos genes *GDF9*, *BMPRI-b* e *BMP15* na População1 (n=75).

Gene	Região do gene										Total	
	†Tamanho	Posição (pb)	*Exons	Referência GB	φ percentual de dados perdidos (n=75)	Transição	Transversão	Indels	Exon (sinônimos)	Exon (não-sinônimos)		Introns
<i>GDF9</i>	4544pb	686 a 5230pb	2	AF078545	0	33	9	6	8	5	19	48
<i>BMPRI-B</i>	54300pb	1538 a 52007pb	10	AF357007.1	1,6%	880	346	235	11	5	1314	1461
<i>BMP15</i>	8683 pb	946 a 7586 pv	2	JN65567.1	0	59	35	7	1	6	69	101
TOTAL						972	390	248	20	16	1402	1610

φ Percentual de SNPs não genotipados que estavam presentes em (1) animais que possuíam regiões com gaps e de desconhecida região sequenciada montada; (2) animais que foram excluídos da re montagem das sequências com *gaps* reduzidos; *número de exons por gene. † Comprimento em pares de bases da região total de cada gene analisado,

Um total de 684 (42%) sítios polimórficos, sendo dez deles localizados em regiões de exon dos genes, foram encontrados entre as raças brasileiras (N=6): Crioula, Morada Nova e Santa Inês, nos genes *GDF9* (9), *BMP15* (18) e *BMPRI-b* (657).

Desvios de EHW foram observados em 40% (634) do total (1610) dos SNPs na Pop1 (N=75), dentre eles, 145 apresentaram valores significativos ($P < 0,05$), 128 ($P < 0,01$) e 361 ($P < 0,001$) (Tabela 4).

Tabela 4– Sítios polimórficos (36) encontrados nas regiões de exon dos genes: *GDF9*, *BMPRI-b* e *BMP15* na Pop1 (n=75).

Genes	SNP	Alteração no mRNA	Exon	Alteração na Proteína.	Qualidade (Phred)	MAF	HetObs.	EHW (P.value)	Referência Bibliográfica.	
<i>GDF9</i>	10G	238c>t	1	Q80Stop	71.3	0,007	0,013	1	-	
	11G	260g>a	1	R87H	999	0,06	0,067	0,0329	Hanrahan et al. 2004	
	12G	282g>a	1	K94K	94.5	0,027	0,027	0,0803	-	
	32G	471c>t	2	V157V	999	0,06	0,067	0,0329	Hanrahan et al. 2004	
	33G	477g>a	2	L159L	999	0,45	0,347	0,0122	Hanrahan et al. 2004/ Chu et al 2011	
	34G	531c>t	2	N177N	6.46	0,007	0	0,0134	-	
	35G	721g>a	2	E241K	999	0,15	0,093	5,22E-02	Hanrahan et al. 2004	
	36G	750a>g	2	L249L	999	0,053	0,067	0,0041	Silva et al. 2011	
	37G	894a>c	2	L298L	999	0,067	0	2,95E-04	-	
	38G	978a>g	2	E326E	999	0,173	0,24	0,2824	Hanrahan et al. 2004	
	39G	994g>a	2	V332I	999	0,17	0,227	0,1024	Hanrahan et al. 2004	
	40G	1203g>a	2	V401V	91.3	0,007	0,013	1	-	
	41G	1357c>t	2	R453C	10.7	0,007	0,013	1	-	
	35Br	1470c>a	10	T490T	999	0,28	0,027	1	Chu et al. 2011	
	57Br	1366c>a	9	R456R	33.6	0,007	0,013	1	-	
	58Br	1299t>c	9	S433S	999	0,05	0,027	3.0E-4	-	
	<i>BMPRI-B</i>	145Br	1147t>c	8	L382L	19	0,008	0,033	1.0	Chu et al. 2011
		146Br	1113c>a	8	R371R	999	0,38	0,307	0,0044	Chu et al. 2011
439Br		1023g>a	7	K341K	999	0,025	0,017	0,0513	-	
440Br		1019t>c	7	V340A	3.89	0,008	0,017	1	-	
441Br		879c>t	7	S293S	19.8	0,008	0,017	1	-	
442Br		864t>c	7	Y288Y	999	0,28	0,356	0.5125	Chu et al. 2011	
488Br		746a>g	6	G249R	999	0,053	0,107	1	Mulsant et al., 2001; Wilson et al., 2001	
489Br		597a>g	6	T199T	999	0,449	0,39	0,1499	-	
490Br		594g>a	6	L195L	999	0,085	0,136	0.6797	-	
737Br		511c>t	5	P171S	8.03	0,007	0,013	1	-	
1088Br		192g>a	2	M64I	999	0,033	0,067	1	Chu et al. 2011	
1363Br		126g>a	1	S42S	999	0,06	0,013	4,16E-02	-	
1364Br		104a>g	1	H335R	13	0,008	0,033	1	-	
11B		28del30/exon1	1	Leu10del	999	0,46	0,28	2.0E-4	Hanrahan et al. 2004	
12B		28del30/exon1	1	Leu10del	999	0,46	0,28	2.0E-4	Hanrahan et al. 2004	
<i>BMP15</i>		13B	116c>g	1	P39R	999	0,03	0,067	1	-
		83B	529c>g	2	P177A	41.4	0,007	0,013	1	-
		84B	747t>c/exon2	2	P249P	20	0,013	0,027	1	Hanrahan et al. 2004
	85B	755t>c	2	L252P	999	0,1	0,12	0,0389	-	
	86B	1037g>a	2	S346N	999	0,05	0,04	0,0095	-	

Tabela 6- Sítios polimórficos encontrados no exon do gene *BMPRI-b* (16) e sua frequência absoluta por raça.

Raças	Animais diplóides (n)	35Br	57Br	58Br	145Br	146Br	439Br	440Br	441Br	442Br	488Br	489Br	490Br	737Br	1088Br	1363Br	1364Br
African White Dorper	2																
Afrikaner, Namaqua	1	1				1											
Afrikaner, Ronderib	2	3										2	1				
Afshari	2	4				1				1		1	1		1		
Awassi	1	1				1											
Awassi, Turkish	2	4								1		1					
Bangladeshi	2					4					2	2					
Brazilian Creole	2	3				1				3		1					
Castellana	2	1				4				2		1					
Changthangi	2	1				1						1	2				
Cheviot	2			2		1				1							2
Churra	2			1			2										3
Cine Capari	1	1										2			1		
Dorset, Poll	1	1				1				1		1					
Ethiopian Menz	1	1										2					
Finnsheep	2									1		3					
Garole	1					2					1	2					
Garole, Indian	1					1					1	2					
Garut	2	1				2			1		2	4					
Gulf Coast native	2					1				4							
Karakas	2									1		1	2				
Karya	1				1	1		1		1							
Lacaune, Meat	1					2				1		1					
Lacaune, Milk	1	1				2											
Merino	3	1				2				1		2					
Morada Nova	2					4						4					1
Norduz	2	3				1						2			1		
Ojalada	2					1				2							
<i>Ovis canadensis</i>	3		1	2		1	1			2						2	
<i>Ovis dalli</i>	2	1				2				1		1		1			
Romney	1	1				2				1		1					
Sakiz	2	2								2		2			2		
Salz	3					1						4					
Santa Inês	2					4				2		2					
Scottish Blackface	1									2							
Sumatra	2	1				3					2	2					
Swiss Mirror	1	2				2				1		2					
Swiss White Alpine	4	3		2		2						1	2			2	
Texel	1	2				2						1					
Tibetan, Eastern	1	1															
Tibetan, Northern	1	1								1			1				
Valais Blacknose	1	1				1						1					
Welsh Hardy Speckled Face	1					2											
Welsh Mountain, Dollgellau	1					1											
Welsh mountain, Tregaron	1									1		1	1				

Tabela 7- Sítios polimórficos encontrados no exon do gene *BMP15* (7) e sua frequência absoluta por raça.

Raças	Animais diplóides (n)	11B	12B	13B	83B	84B	85B	86B
African White Dorper	2	3	1				1	
Afrikaner, Namaqua	1	2					1	
Afrikaner, Ronderib	2	4					2	
Afshari	2		4					
Awassi	1		2					
Awassi, Turkish	2		4					
Bangladeshi	2	4					1	
Brazilian Creole	2	2	1					
Castellana	2		4					
Changthangi	2		3		1		1	
Cheviot	2	2	1	1				1
Churra	2	2	2	2				3
Cine Capari	1	2						
Dorset, Poll	1		2					
Ethiopian Menz	1	1						
Finnsheep	2	4						
Garole	1	1	1				1	
Garole, Indian	1	1	1					
Garut	2	4					3	
Gulf Coast native	2	1	2					
Karakas	2		4					
Karya	1		1					
Lacaune, Meat	1	2						
Lacaune, Milk	1		2					
Merino	3	5	1					
Morada Nova	2	2	2					
Norduz	2	1	2					
Ojalada	2	1	3			1		
<i>Ovis canadensis</i>	3	4	1	1		1		1
<i>Ovis dalli</i>	2	1	3					
Romney	1		2					
Sakiz	2		4					
Salz	3	6						
Santa Inês	2	3						
Scottish Blackface	1	1						
Sumatra	2	3	1				3	
Swiss Mirror	1		2					
Swiss White Alpine	4	4	2	1				2
Texel	1		2					
Tibetan, Eastern	1		1					
Tibetan, Northern	1		2					
Valais Blacknose	1	2						
Welsh Hardy Speckled Face	1	1	1				1	
Welsh Mountain, Dollgellau	1		2					
Welsh mountain, Tregaron	1		2					

3.3 Estimativas de haplótipos formados pelos SNPs identificados nos genes *GDF9*, *BMP15* e *BMPRI-b*.

Para os três genes separadamente foram estimados, a partir de 1610 SNPs, 45 haplótipos para *GDF9*, 95 para *BMP15* e 150 para *BMPRI-b*, com diversidade haplotípica (DH) igual a 0,807, 0,98 e 1, respectivamente.

No gene *GDF9*, os haplótipos 2 e 7 (HG2 e HG7) se destacam pelas altas frequências: 0,407 (N = 61, em 29 raças) e 0,16 (N = 24, em 15 raças), respectivamente, e não foram observados SNPs com alteração não conservativas nesses haplótipos. Treze haplótipos apresentaram SNPs com mutações funcionais em 15 diferentes raças (Tabela 1A - Anexo), destacando-se os haplótipos HG1, HG14 e HG45 observados em mais de uma raça. O HG1 foi observado em cinco raças (Changthangi, Morada Nova, Namaqua Afrikaner, Pool Dorset e Ronderib Afrikaner), o HG14 em quatro raças (Castellana, Karakas, Norduz e *O. Canadensis*) e o HG45 em duas raças (Changthangi e Cheviot) (Tabela 1A - Anexo). Não foi observada a presença de dois ou mais alelos que causam mutações não conservativas no mesmo haplótipo.

Para o gene *BMP15*, 20 haplótipos estão presentes em duas ou mais raças, sendo o haplótipo HB46 o mais frequente na população, 0,07 (N = 11, em nove raças) sem SNPs com alterações não conservativas nos AAs. Oito haplótipos apresentaram mutações não conservativas, contendo o SNP 85B e InDel 11-12B, estes foram observados em mais de uma raça, destacando os haplótipos HB3 e HB8, observados em oito e quatro raças diferentes, respectivamente (Tabela 1A - Anexo). Os seis haplótipos restantes foram observados no máximo em duas raças cada um.

Haplótipos com SNPs que apresentavam mutações não conservativas foram observados em 23 raças da Pop1 (Tabela 1A - Anexo). Para esse gene a frequência de dois alelos que causam alterações não conservativas, representados pelos SNPs 13B e 86B, foram encontrados na mesma fase haplotípica. Esses alelos estão presentes em três haplótipos nas raças: Churra (HB70, HB76) e Swiss White Alpine (HB65) (Tabela 1A - Anexo). Outros sete haplótipos também possuem a combinação da deleção 11-12B e o alelo 85B, os quais alteram sequências de aminoácidos na proteína de *BMP15*. Esses sete haplótipos estão distribuídos nas raças: Garut (HB14), Namaqua Afrikaner (HB18), Sumatra (HB48 e HB49), Ronderib Afrikaner (HB42 e HB43) (Tabela 1A - Anexo).

Para o gene *BMPRI-b* o valor da DH=1, com 150 haplótipos estimados, dois haplótipos para cada indivíduo. Essa diversidade provavelmente é explicada pela extensão da

região gênica analisada (54300 pb), além da alta diversidade entre indivíduos da população estudada. Nesse gene também não foi observada a frequência de dois ou mais alelos que causam mutações não conservativas no mesmo haplótipo (Tabela 1A - Anexo).

O valor da diversidade haplotípica entre os três genes foi maior (proporcionalmente) de acordo com o tamanho (em pb) de cada região gênica analisada. A alta diversidade haplotípica ($DH > 0,8$) encontrada em cada gene demonstra a grande diversidade genética entre os animais e as raças da Pop1.

A estimativa de haplótipos para os três genes juntos, a partir de 1610 SNPs, resultou na formação de 150 haplótipos ($N = 75$), com DH igual a 1, ou seja, alta diversidade haplotípica na população, onde cada indivíduo mostrou um haplótipo único.

Adicionalmente foram estimados novos haplótipos utilizando somente os SNPs (14) e a InDel (1) (Tabela 8) representando as mutações não conservativas na região codante dos três genes em conjunto. Foram estimados 24 haplótipos nas 150 sequências (Tabela 8).

Os haplótipos foram nomeados como HEx (haplótipos em exon) de 1 a 24 para cada haplótipo específico. O haplótipo HEx1 foi o mais frequente ($f=0,28$, $N=42$) e apresentou todos os alelos de cada polimorfismo igual à sequência referência do genoma (REFOav3.1-<http://www.livestockgenomics.csiro.au/sheep/oar3.1.php>). Nos haplótipos onde se observou variações em relação à sequência referência do genoma, cinco (HEx2, HEx4, HEx5, HEx8 e HEx11) foram representados por apenas um SNP e 20 haplótipos com dois ou mais SNPs (de 2 a 4) que causaram alterações não conservativas nos aminoácidos (Tabela 8).

No total, nessa segunda estimativa, 16 haplótipos apresentaram duas ou mais variantes em dois genes diferentes, sendo três haplótipos com variação nos genes *GDF9* e *BMPRI-b*, nove em *GDF9* e *BMP15* e quatro em *BMP15* e *BMPRI-b* (Tabela 8). O InDel 11-12B do gene *BMP15* é encontrado em 15 haplótipos, sendo 12 destes com pelo menos um SNP nos genes *GDF9* e *BMPRI-b*. A variante denominada Booroola (*FecB^B*), foi encontrada em dois haplótipos (HEx6 e HEx7) e observada sempre com alguma das variantes do gene *BMP15* (Tabela 8).

Tabela 8- Segunda estimativa de haplótipos utilizando SNPs encontrados nos exons dos genes *GDF9*, *BMPRI-b* e *BMP15* (com mutação não conservativa) e suas respectivas mutações e frequências.

Haplótipo	Frequência (N)*	<i>GDF9</i> (CR.5)					<i>BMPRI-B</i> (CR.6)					<i>BMP15</i> (CR.X)				
		10G	11G	35G	39G	41G	440Br	488Br	737Br	1088Br	1364Br	11-12B	13B	83B	85B	86B
HEx.1	0,28 (42)	C	G	G	G	C	A	T	G	C	T	-	C	C	T	G
HEx.2	0,24 (36)	CTT	
HEx.3	0,04 (6)	CTT	.	.	C	.	
HEx.4	0,026 (4)	T	
HEx.5	0,026 (4)	.	.	A	
HEx.6	0,05 (7)	C	.	.	.	CTT	.	.	C	.	
HEx.7	0,006 (1)	C	.	.	.	CTT	
HEx.8	0,05 (7)	.	.	.	A	
HEx.9	0,1 (15)	.	.	.	A	CTT	
HEx.10	0,05 (7)	.	A	A	
HEx.11	0,006 (1)	G	.	.	
HEx.12	0,01 (2)	.	.	A	CTT	
HEx.13	0,033 (5)	.	.	A	CTT	G	.	.	A	
HEx.14	0,006 (1)	T	CTT	
HEx.15	0,006 (1)	.	.	A	A	
HEx.16	0,006 (1)	T	.	CTT	
HEx.17	0,006 (1)	T	CTT	
HEx.18	0,006 (1)	.	A	A	.	.	G	
HEx.19	0,006 (1)	C	CTT	
HEx.20	0,006 (1)	.	.	.	A	C	
HEx.21	0,006 (1)	.	.	.	A	.	.	A	
HEx.22	0,01 (2)	.	.	.	A	CTT	.	.	C	.	
HEx.23	0,01 (2)	.	A	A	CTT	
HEx.24	0,006 (1)	.	.	A	A	

**N= Número absoluto de haplótipos.

No total 436 amostras foram amplificadas e sequenciadas na região do segundo exon do gene *GDF9*, sendo, 186 amostras da raça Santa Inês (OSI) e 253 da raça Morada Nova (OMN). A amplificação gerada via PCR resultou em diferentes tamanhos de fragmentos sequenciados em cada indivíduo que variaram entre 270 a 392 pb e foram alinhadas entre a posição 3746 a 4137 pb da sequência referência depositada no *GeneBank* (AF078545). Como o tamanho do fragmento variou por indivíduos, foi necessário padronizar o tamanho dos fragmentos de todos os indivíduos reduzindo para 270 pb de sequência em comum em toda amostragem usados para a busca de polimorfismo e estimativa de haplótipos.

Foram identificados 12 SNPs no exon 2 do gene *GDF9*. Os SNPs foram nomeados de 1 a 12 sequencialmente de acordo com sua posição em relação à sequência referência usada para o alinhamento (AF078545), além da identificação da letra G (referente ao gene *GDF9*) e B (referente a População 2 ser composta de raças brasileiras) (Tabela 9). Destes 12 SNP encontrados, sete foram alterações do tipo transversões e cinco transições, e sete SNPs resultaram em alteração não conservativa em aminoácidos (Tabela 8). Quatro polimorfismos, (2GB, 9GB, 10GB, e 11GB) já haviam sido relatados na literatura nas raças Belclare/Cambridge e Santa Inês (Hanrahan et al., 2004 e Silva et al., 2011).

O número médio de alelos por SNP foi de 2,08, com apenas um SNP (2GB) apresentando três alelos (G, T C), que também resultaram em diferentes AA na região 291 da proteína, de G291C (alelo T) e G291R (alelo C) (Tabela 9). Quatro SNPs apresentaram MAF baixo ($MAF < 0,01$) (Tabela 9), no entanto, o índice de qualidade Phred (QV) variou de 14 a 30, para os mesmos; 5GB (QV= 14 a 20), 6GB (QV=21 a 30), 7GB (QV=14 a 20), e 8GB (QV=22 a 29), assumindo a confiabilidade da existência destes polimorfismos.

Todos os polimorfismos encontrados estavam presentes em uma ou nas duas raças. O alelo mais frequente em toda população foi o alelo G (0,56), do SNP 11GB, presente nas duas raças da Pop2 (Tabela 9).

Desvios de EHW foram observados para os SNPs: 1GB ($P < 0,001$), 7GB ($P < 0,001$) e 12GB ($P < 0,01$); quando analisados na população total ($N=436$). As análises de EHW por raça mostraram que os SNPs 1GB ($P < 0,05$) e 11GB ($P < 0,05$) apresentam desvios de EHW na raça OMN; e os SNPs 1GB ($P < 0,001$), 7GB ($P < 0,001$) e 12GB ($P < 0,001$) na raça OSI.

Tabela 9- Caracterização dos SNPs encontrados via re-sequenciamento do exon 2 do gene *GDF9* na Pop2.

SNP	Alteração	Posição (pb)		Alteração no AA	Número de Alelos	MAF (Frequência do alelo Menor)			HetObs.
		Cr. 5 REFOav3. 1*	mRNA			Total (n=43 6)	OMN ¹ (n=25 3)	OSI ¹ (n=18 6)	
1GB	A>T	41841537	859a>t	K287stop	2	0,213	0,156	0,271	0,43
2GB	G>T/C	41841525	871a>t	G291C /G291R	3	0,015	0,019	0,009	0,03
3GB	G>C	41841505	891g>c	G297G	2	0,035	0,031	0,042	0,07
4GB	T>C	41841466	930t>c	G310G	2	0,031	0,026	0,039	0,06
5GB	G>A	41841457	939g>a	S313S	2	0,008	0,01	0,005	0,02
6GB	C>G	41841446	951c>g	R317R	2	0,005	0,004	0,005	0,01
7GB	G>C	41841444	953g>c	R318T	2	0,008	0,004	0,014	0,01
8GB	G>C	41841442	955g>c	D319H	2	0,003	0,002	0,005	0,01
9GB	A>G	41841418	978g>a	E326E	2	0,031	0,03	0,03	0,06
10GB	G>A	41841402	994g>a	V332I	2	0,034	0,032	0,035	0,07
11GB	T>G	41841362	1034t>g	F345C	2	0,44	0,44	0,44	0,46
12GB	T>A	41841182	1214t>a	I405N	2	0,143	0	0,325	0,29

1 – Raças brasileiras utilizadas: OMN – Ovino Morada Nova, OSI – Ovino Santa Inês

* <http://www.livestockgenomics.csiro.au/sheep/oar3.1.php>

3.4 Estimativas de haplótipos no exon 2 do gene *GDF9* para População2.

As análises para gerar uma estimativa de haplótipos foram realizadas utilizando os doze SNPs encontrados nessa população (Tabela 10); obtendo-se 30 haplótipos únicos com DH= 0,73, compreendendo uma região de 355pb entre as posições 3763 a 4118pb da sequência de referência (AF078545) do gene *GDF9*.

O haplótipo denominado 1 (HGB1) foi o mais frequente na população (0,4), seguido do haplótipo 2 (HGB2 –f=0,3) que é idêntico ao haplótipo presente na sequência de referência (Tabela 10). A diferença entre o HGB1 e a sequência referência, é apenas por um SNP (11GB). Esse SNP é um polimorfismo altamente frequente em populações da raça brasileira de ovelhas prolíficas Santa Inês (Silva et al. 2011).

Tabela 10- Haplótipos estimados na Pop2 (N=436,) na região do exon 2 do gene *GDF9*.

Haplótipos da Pop2 (N=872) *	Frequência (n)	1GB	2GB	3GB	4GB	5GB	6GB	7GB	8GB	9GB	10GB	11GB	12GB
		A	G	G	T	G	C	G	G	A	G	T	T
HGB1	0,40 (358)	G	.
HGB2	0,30 (266)
HGB3	0,02 (19)	G	A	.	.	A
HGB4	0,22 (84)	T	G	.
HGB5	0,006 (5)	.	C	A
HGB6	0,006 (5)	G	A	.	.	.
HGB7	0,004 (4)	A	G	.
HGB8	0,007 (6)	T
HGB9	0,011 (10)	T	.	C	C
HGB10	0,004 (4)	A
HGB11	0,007 (6)	T	T	C	C
HGB12	0,002 (2)	.	.	C
HGB13	0,002 (2)	G	C	.	.	.	G	.
HGB14	0,001 (1)	T	T	C
HGB15	0,009 (8)	T	.	C	G	.
HGB16	0,001 (1)	N	N	N	.	A	.	C	C	.	.	G	.
HGB17	0,001 (1)	.	.	C	C
HGB18	0,007 (6)	T	.	.	C
HGB19	0,002 (2)	T	.	C	C	G	.
HGB20	0,001 (1)	T	.	.	C	G	.
HGB21	0,001 (1)	.	.	.	C
HGB22	0,002 (2)	C
HGB23	0,001 (1)	A	.	C	C	.	.	G	.
HGB24	0,001 (1)	A	G	C	C	.	.	G	.
HGB25	0,008 (7)	T	A
HGB26	0,047(41)	A
HGB27	0,014 (12)	T	G	A
HGB28	0,022 (20)	G	A
HGB29	0,001 (1)	G	A	G	.	.
HGB30	0,001 (1)	G	.	.	.	A

*N= Número de haplótipos.

4 DISCUSSÃO

Os SNPs são considerados unidades fundamentais de variações genéticas e, atualmente dentre os marcadores genéticos é o mais indicado para compor painéis de genotipagem por serem abundantes em todo o genoma (Kijas et al., 2012), e possuírem grande potencial para automação e baixos custos em plataformas de genotipagens (Ding & Jin, 2009).

Atualmente, estudos de busca por variantes genéticas em raças prolíficas de ovinos tem sido realizadas principalmente com objetivo de se detectar SNPs de grande efeito com características de prolificidade (Chu et al., 2011a; López-Ramírez et al., 2014; Mullen et al., 2013; Souza et al., 2014).

Este estudo descreveu a identificação de 1610 SNPs em genes relacionados à prolificidade em ovinos baseado em sequenciamento de genes completos e de 12 SNPs baseados em fragmentos do exon 2 do gene *GDF9*. Dos SNPs identificados na Pop 1 (ISGC, N=75), 36 estão localizados em exons nos genes: *GDF9* (13), *BMP15* (7) e *BMPRI-b* (16).

Para o gene *GDF9*, oito dos 13 SNPs em exon já foram identificados previamente por outros estudos. Os SNPs: 11G, 32G, 33G, 35G, 38G e 39G já foram descritos nas raças Belclare e Cambridge (Hanrahan et al., 2004), o 36G em ovinos da raça Santa Inês (Silva et al., 2011), e mais recentemente, o SNP 35G foi também encontrado no mRNA do gene *GDF9* em ovários de ovelhas da raça Afshari (Eghbalsaied et al., 2012). Em nosso trabalho o SNP 35G também foi observado em indivíduos da raça Afshari, ($f=0,15$) assim como nas raças Awassi, Castellana, Cheviot, Churra, Karakas, Karya, Norduz, Ojalada, *O. canadensis*, Salz e Swiss White Alpine.

O InDel 11-12B, e o SNP 84B, presentes no gene *BMP15*, também foram identificados previamente por Hanrahan et al. 2004 nas raças Belclare e Cambridge, sem nenhum efeito associado com características de prolificidade. No trabalho atual os SNPs 13B e 86B, com variantes que provocam alterações não conservativas nos AAs, foram relatados pela

primeira vez em quatro animais de raças consideradas prolíficas: Cheviot (1 animal), Churra (2 animais) e Swiss White Alpine (1 animal). Os SNPs13B e 86B por serem mutações com alterações não conservativas no aminoácidos e presentes em exon de raças prolíficas serão importantes candidatos para compor o painel de genotipagem associado a prolificidade.

No gene *BMPRI-b*, o SNP 488Br, foi encontrado no presente estudo nas raças Garole, Garut, Sumatra e Bangladeshi (Tabela 6) e anteriormente identificado por vários autores em diferentes raças, sendo primeiramente reportado na raça Merino (Davis et al., 1982, Mulsant et al., 2001; Wilson et al., 2001) e depois nas raças Chinesas Hu (Davis et al., 2006), e Small Tain Han (Chu et al., 2011), Garole (Polley et al., 2010), Javanese (Davis et al., 2002) e Bonpala (Roy et al., 2011). Em nosso estudo a presença do SNP 488Br foi identificado nas raças prolíficas (pela primeira vez nessas raças) de origem asiáticas: Bangladeshi da Índia (Pervage et al., 2012), Garut e Sumatra da Indonésia (Mason, 1980). A raça Garut surgiu de cruzamentos da raça Merino (Mason, 1996), o que portanto sugere que essa variante pode ter sido herdada a partir desta raça.

Outras cinco variantes encontradas no gene *BMPRI-b*, 35Br, 145Br, 146Br, 442Br e 1088Br também já foram identificadas por Chu et al. (2011) nas raças chinesas prolíficas de ovinos Small Tail Han e Hu (Tabela 5), mas sem nenhum efeito conhecido associado com a prolificidade.

Os 634 SNPs identificados com desvios de EHW ($0,05 > P < 0,001$) na Pop1 podem ter se apresentado fora do equilíbrio pelo fato dos mesmos estarem sendo afetados pela seleção artificial no processo de domesticação dos ovinos, já que a Pop1 tem alta diversidade entre os animais das diferentes raças do mundo selecionados para representar a diversidade da espécie ovina (Kijas et al., 2012). A presença de 1404 SNPs com $MAF > 0,01$ nos três genes evidencia alta diversidade das amostras e a alta diversidade gênica dos genes relacionados a taxa de ovulação em ovinos.

Seis SNPs (1GB, 2GB, 3GB, 4GB, 11GB e 12GB) identificados no segundo exon do gene *GDF9*, na Pop2 com $MAF > 0,015$, sugerem ser específicos de raças brasileiras, já que foram observados apenas nas raças brasileiras localmente adaptadas: Morada Nova e Santa Inês. Adicionalmente, os SNPs 2GB e 11GB já foram previamente identificados por Silva et al. (2011) em ovelhas prolíficas da raça Santa Inês.

Baseado nas estimativas haplotípicas utilizando SNPs com mutações não conservativas, foram observados haplótipos com presença de duas ou mais variantes alélicas entre genes diferentes. A frequência de haplótipos ($f=0,2$) com mutações nos genes *GDF9* e *BMP15* varia de acordo com as raças, uma vez que, estas mutações foram encontradas em nove

haplótipos, que diferiram entre cinco SNPs em *GDF9* (10G, 11G, 35G, 39G e 41G) e quatro em *BMP15* (InDel 11-12B, 13B, 85B e 86B). Haplótipos com mutações nos genes *BMPRI-b* e *GDF9* (4 haplótipos, $f=0,066$) e *BMPRI-B* e *BMP15* (3 haplótipos, $f=0,02$) foram também observados na Pop1. Destaca-se o haplótipo Hex13 ($f=0,006$), representado por quatro SNPs com variantes que causam alterações não conservativas nos AAs dos genes *GDF9* (35G) e *BMP15* (InDel 11-12B, 13B e 86B). Esse haplótipo (Hex13) esteve presente nos mesmos animais citados anteriormente das raças Churra, Cheviot e Swiss Alpine.

Os resultados de frequência de haplótipos representados por uma ou mais mutações que alteram aminoácidos das proteínas em um ou dois genes diferentes, podem sugerir a associação dos genes relacionados à taxa de ovulação em ovinos. Esses resultados juntos com estudos genéticos encontrados na literatura (Hanrahan et al., 2004; Barzegari et al., 2009; Chu et al., 2007; Polley et al., 2010) sugerem que as variantes em genes que afetam a taxa de ovulação podem estar ligadas entre si, além de terem efeitos variados nos fenótipos de diferentes raças.

Como resultado deste estudo, foram identificados, a partir da mineração nos genes *GDF9*, *BMP15* e *BMPRI-b*, polimorfismos que poderão compor um painel reduzido e informativo de SNPs relacionados a características de prolificidade em ovinos. A validação destes SNPs, e a associação de um painel reduzido composto pelos mais informativos, com características de prolificidade em ovinos, devem ser realizadas em análises futuras em raças prolíficas com histórico de parto simples e múltiplos. Os polimorfismos, depois de validados em populações de raças prolíficas, poderão auxiliar programas de conservação e melhoramento de raças de ovinos ao redor do mundo.

5 CONCLUSÃO

A mineração das sequências completas dos genes *GDF9*, *BMP15* e *BMPRI-b* utilizando a População 1 identificou um total de 1610 polimorfismos nas diferentes raças de ovinos do mundo.

Os SNPs 488Br, 13B e 86B identificados na região de exon dos genes *BMPRI-b* e *BMP15*, que resultam em alterações não conservativas nos aminoácidos, são candidatos imediatos a serem testados nas raças conhecidas por serem prolíficas.

O SNP 11GB foi encontrado apenas na Pop2, de raças brasileiras. Esse SNP é um polimorfismo altamente frequente em populações da raça brasileira deslanadas de ovelhas prolíficas.

6 REFERÊNCIAS BIBLIOGRÁFICAS

- ANDERSSON, L.; GEORGES, M. Domestic-animal genomics: deciphering the genetics of complex traits. **Nature Reviews Genetics**, v. 5, n. 3, p. 202–212, 2004.
- ANSORGE, W. J. Next-generation DNA sequencing techniques. **New biotechnology**, v. 25, n. 4, p. 195–203, abr. 2009.
- BARCHI, L. et al. Identification of SNP and SSR markers in eggplant using RAD tag sequencing. **BMC Genomics**, v. 12, n. 1, p. 304, 2011.
- BARRIS, W. et al. Next generation sequencing of African and Indicine cattle to identify single nucleotide polymorphisms. **Animal Production Science**, v. 52, n. 2-3, p. 133–142, 2012.
- BARZEGARI, A. et al. Polymorphisms in *GDF9* and *BMP15* Associated with Fertility and Ovulation Rate in *Moghani* and *Ghezel* Sheep in Iran. **Reproduction in Domestic Animals**, v. 669, p. 666–669, 2009.
- BENNETT, S. Solexa Ltd. **Pharmacogenomics**, v. 5, n. 4, p. 433–438, 2004.
- BODENSTEINER, K. J. et al. Molecular cloning of the ovine Growth/Differentiation factor-9 gene and expression of growth/differentiation factor-9 in ovine and bovine ovaries. **Biology of reproduction**, v. 60, n. 2, p. 381–386, 1999.
- BOREVITZ, J. O.; CHORY, J. **Genomics tools for QTL analysis and gene discovery** *Current Opinion in Plant Biology*, 2004.
- CALVO, J. H. et al. Isolation, mapping and identification of SNPs for four genes (*ACP6*, *CGN*, *ANXA9*, *SLC27A3*) from a bovine QTL region on BTA3. **Cytogenetic and Genome Research**, v. 114, n. 1, p. 39–43, 2006.
- CHU, M. et al. Polymorphisms of coding region of *BMPR-IB* gene and their relationship with litter size in sheep. **Mol Biol Rep**, v. 38, n. 6, p. 4071–4076, 2011.
- CHU, M. X. et al. Mutations in *BMPR-IB* and *BMP-15* genes are associated with litter size in Small Tailed Han sheep (*Ovis aries*). **J Anim Sci**, v. 85, n. 3, p. 598–603, 2007.
- DAS, S. et al. Deploying QTL-seq for rapid delineation of a potential candidate gene underlying major trait-associated QTL in chickpea. **DNA Research**, v. 22, n. 3, p. 193–203, 2014.
- DAVIS, G. H. et al. **Segregation of a major gene influencing fecundity in progeny of Booroola sheep** *New Zealand Journal of Agricultural Research*, 1982. Disponível em: <<http://dx.doi.org/10.1080/00288233.1982.10425216>>
- DAVIS, G. H. et al. DNA tests in prolific sheep from eight countries provide new evidence on origin of the Booroola (*FecB*) mutation. **Biol Reprod**, v. 66, n. 6, p. 1869–1874, 2002.
- DAVIS, G. H. et al. Investigation of the Booroola (*FecB*) and Inverdale (*FecX(I)*) mutations in 21 prolific breeds and strains of sheep sampled in 13 countries. **Anim Reprod Sci**, v. 92, n. 1-2, p. 87–96, 2006.
- DE CAMARGO, G. M. F. et al. First polymorphisms in *JY-1* gene in cattle (*Bos taurus indicus*) and their association with sexual precocity and growth traits. **Molecular Biology Reports**, p. 1–5, 2012.

- DING, C.; JIN, S. High-throughput methods for SNP genotyping. **Methods in Molecular Biology**, v. 578, p. 245–254, 2009.
- EGHBALSAIED, S. et al. Presence of SNPs in GDF9 mRNA of Iranian Afshari Sheep. v. 5, n. 4, p. 225–230, 2012.
- ERNST, C. W.; STEIBEL, J. P. **Molecular advances in QTL discovery and application in pig breeding** *Trends in Genetics*, 2013.
- FACÓ, O. et al. **Teste de Desempenho Individual de Reprodutores da Raça Morada Nova: Resultados da Prova em Morada Nova - CE - 18/02 a 04/06/2008**. Sobral: [s.n.]. Disponível em: <<http://www.cnpc.embrapa.br>>.
- FAHRENKRUG, S. C. et al. Single nucleotide polymorphism (SNP) discovery in porcine expressed genes. **Animal Genetics**, v. 33, n. 3, p. 186–195, 2002.
- FONTANESI, L. et al. Next Generation Semiconductor Based-Sequencing of a Nutrigenetics Target Gene (GPR120) and Association with Growth Rate in Italian Large White Pigs. **Animal biotechnology**, v. 26, n. 2, p. 92–7, 2015.
- GODDARD, M. E.; HAYES, B. J. Mapping genes for complex traits in domestic animals and their use in breeding programmes. **Nature Reviews Genetics**, v. 10, n. 6, p. 381–391, 2009.
- HANRAHAN, J. P. Mutations in the Genes for Oocyte-Derived Growth Factors GDF9 and BMP15 Are Associated with Both Increased Ovulation Rate and Sterility in Cambridge and Belclare Sheep (*Ovis aries*). **Biology of Reproduction**, v. 70, n. 4, p. 900–909, 2004.
- HEATON, M. P. et al. Genetic Testing for TMEM154 Mutations Associated with Lentivirus Susceptibility in Sheep. **PLoS ONE**, v. 8, n. 2, p. e55490, 2013.
- HEATON, M. P. et al. SNPs for parentage testing and traceability in globally diverse breeds of sheep. **PloS one**, v. 9, n. 4, p. e94851, 2014.
- JIANG, Y. et al. The sheep genome illuminates biology of the rumen and lipid metabolism. **Science (New York, N.Y.)**, v. 344, n. 6188, p. 1168–73, 2014.
- KIJAS, J. W. et al. Genome-Wide Analysis of the World's Sheep Breeds Reveals High Levels of Historic Mixture and Strong Recent Selection. **PLoS Biol**, v. 10, n. 2, p. e1001258, 2012.
- KUMAR, S.; YOU, F. M.; CLOUTIER, S. Genome wide SNP discovery in flax through next generation sequencing of reduced representation libraries. **BMC genomics**, v. 13, n. 1, p. 684, 2012.
- LI, H. et al. The Sequence Alignment/Map format and SAMtools. **Bioinformatics**, v. 25, p. 2078–2079, 2009.
- LIBRADO, P.; ROZAS, J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. **Bioinformatics**, v. 25, n. 11, p. 1451–1452, 2009.
- LÓPEZ-RAMÍREZ, R. B. et al. Analysis of the 3' End Regions of the Gdf9 and Bmpr1B Genes in Blackbelly Sheep From Yucatán, Mexico. **Ciencia e investigación agraria**, v. 41, n. 1, p. 23–24, 2014.
- MARGULIES, M. et al. Genome sequencing in microfabricated high-density picolitre reactors. **Nature**, v. 437, n. 7057, p. 376–80, 2005.
- MASON, I. L. **Prolific tropical sheep** *FAO ANIMAL PRODUCTION AND HEALTH PAPER*. Roma: [s.n.]. Disponível em: <<http://www.fao.org/3/a-x6517e/index.html>>.
- MASON, I. L. **A World Dictionary of Livestock Breeds, Types and Varieties**. Fourth Ed. ed. [s.l.] C.A.B International., 1996.
- MEIRA, C. T. et al. Prospection of genomic regions divergently selected in racing line of Quarter Horses in relation to cutting line. p. 1754–1764, 2014.
- MEUWISSEN, T. H. E.; HAYES, B. J.; GODDARD, M. E. Prediction of total genetic value using genome-wide dense marker maps. **Genetics**, v. 157, n. 4, p. 1819–1829, 2001.
- MULLEN, M. P. et al. Investigation of prolific sheep from UK and Ireland for evidence on origin of the mutations in BMP15 (FecX(G), FecX(B)) and GDF9 (FecG(H)) in Belclare and Cambridge sheep. **PloS one**, v. 8, n. 1, p. e53172, 2013.

- MULSANT, P. et al. Mutation in bone morphogenetic protein receptor-IB is associated with increased ovulation rate in Booroola Merino ewes. **Proc Natl Acad Sci U S A**, v. 98, n. 9, p. 5104–5109, 2001.
- PEAKALL, R.; SMOUSE, P. E. GenALEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. **Bioinformatics**, v. 28, n. 19, p. 2537–2539, 2012.
- PERVAGE, S. et al. Phenotypic Characteristics Of Indigenous Sheep Of Bangladesh. **Bangladesh Journal of Animal Science**, v. 38, n. 1-2, p. 1–6, 2012.
- POLLEY, S. et al. Polymorphism of BMPR1B, BMP15 and GDF9 fecundity genes in prolific Garole sheep. **Tropical Animal Health and Production**, v. 42, n. 5, p. 985–993, 2010.
- PURVIS, I. W.; FRANKLIN, I. R. Major genes and QTL influencing wool production and quality: a review. **Genetics, selection, evolution : GSE**, v. 37 Suppl 1, p. S97–S107, 2005.
- ROBINSON, J. T. et al. Integrative genomics viewer. **Nature Biotechnology**, v. 29, n. 1, p. 24–26, 2011.
- ROTHSCHILD, M. F.; SOLLER, M. Candidate gene analysis to detect traits of economic importance in domestic livestock. **Probe**, v. 8, p. 13–21, 1997.
- ROY, J. et al. Polymorphism of fecundity genes (FecB, FecX, and FecG) in the Indian Bonpala sheep. **Animal biotechnology**, v. 22, n. 3, p. 151–162, 2011.
- SAHANA, G. et al. Genome-wide association study using high-density single nucleotide polymorphism arrays and whole-genome sequences for clinical mastitis traits in dairy cattle. **Journal of dairy science**, v. 97, n. 11, p. 7258–75, 2014.
- SCHNEIDER, S.; ROESSLI, D.; EXCOFFIER, L. Arlequin, Version 2.0: A software for population genetics data analysis. **A software for Population Genetics Data ...**, p. 111, 2000.
- SILVA, B. D. M. et al. A new polymorphism in the Growth and Differentiation Factor 9 (GDF9) gene is associated with increased ovulation rate and prolificacy in homozygous sheep. **Animal Genetics**, v. 42, n. 1, p. 89–92, 2011.
- SOUZA, C. J. H. et al. Mutation in the protease cleavage site of GDF9 increases ovulation rate and litter size in heterozygous ewes and causes infertility in homozygous ewes. **Animal genetics**, v. 45, n. 5, p. 732–9, 2014.
- TAMURA, K. et al. MEGA6: Molecular evolutionary genetics analysis version 6.0. **Molecular Biology and Evolution**, v. 30, n. 12, p. 2725–2729, 2013.
- VALOUEV, A. et al. A high-resolution, nucleosome position map of *C. elegans* reveals a lack of universal sequence-dictated positioning. **Genome research**, v. 18, n. 7, p. 1051–63, 2008.
- VENERONI, G. B. et al. Prospecting candidate snps for backfat in canchim beef cattle. **Genetics and Molecular Research**, v. 9, n. 4, p. 1997–2003, 2010.
- WILSON, T. et al. Highly prolific Booroola sheep have a mutation in the intracellular kinase domain of bone morphogenetic protein IB receptor (ALK-6) that is expressed in both oocytes and granulosa cells. **Biology of reproduction**, v. 64, n. 4, p. 1225–1235, 2001.
- ZERBINO, D. R.; BIRNEY, E. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. **Genome Research**, v. 18, n. 5, p. 821–829, 2008.
- ZHANG, Z. et al. A greedy algorithm for aligning DNA sequences. **Journal of computational biology : a journal of computational molecular cell biology**, v. 7, n. 1-2, p. 203–214, 2000.
- ZOU, X. et al. Genome-wide single nucleotide polymorphism and Insertion-Deletion discovery through next-generation sequencing of reduced representation libraries in common bean. **Molecular Breeding**, v. 33, n. 4, p. 769–778, 2014.

CAPITULO 4

CONSIDERAÇÕES FINAIS E CONCLUSÕES

Os principais resultados desse trabalho mostraram que a estratégia de busca de polimorfismos em genes de interesse econômico cuja característica é controlada por poucos genes tem sua aplicação potencializada, mesmo com a forte tendência da implementação da seleção genômica nos programas de melhoramento tradicionais. Na presente tese, três genes (*GDF9*, *BMP15* e *BMPI1-b*) relacionados com taxa de ovulação/número de cordeiros nascidos por gestação foram escolhidos. A razão principal para essa escolha foi que um processo de seleção, mediado por marcadores moleculares identificados nesses genes, poderá ter um impacto rápido e direto nos rebanhos da espécie ovina.

No Capítulo 2 os objetivos e resultados foram voltados para uma mutação no gene *GDF9* (*FecG^E*) comprovadamente relacionada a um aumento da taxa de ovulação na raça localmente adaptada Santa Inês (Silva et al., 2011). O primeiro resultado identificado foi que essa mutação não é restrita a raça Santa Inês e foi identificada em alta frequência na raça Morada Nova. O segundo resultado mais expressivo foi que esta mutação aparenta não estar associada a características de produção/crescimento reconhecidamente importantes em programas de melhoramento animal avaliadas em teste de desempenho. Não foi observada diferença estatística entre animais com altos e baixos escores em testes de desempenho da raça Morada Nova. Estudos econômicos realizados por Lobo et al.(2011) identificaram que número de borregos nascidos por parto é um importante parâmetro para o lucro das fazendas. Dessa forma, a adição desse teste genético no programa de melhoramento participativo da raça poderá agregar importante informação aos produtores e dar a eles subsídios para as escolhas dos acasalamentos.

O Capítulo 3 teve como foco principal usar as vantagens sobre dados de sequenciamento de nova geração como uma ferramenta para identificação de novos SNPs em

três genes relacionados a taxa de ovulação. Dessa forma, a partir de uma amostragem de 75 animais representantes de várias raças de ovinos espalhadas pelo mundo, foram identificadas novas variantes em genes associados a taxa de ovulação, além de mutações conhecidas já associadas com prolificidade em animais de diferentes raças prolíficas: Morada Nova (*FecGE-Capitulo 2-GDF9*), Bangladeshi (488Br- *BMPBR1-b*) Garut (488Br- *BMPBR1-b*), Garole (488Br- *BMPBR1-b*), Sumatra (488Br - *BMPBR1-b*), Churra (35G- *GDF9*, 13B e 86B- *BMP15*), Cheviot (35G- *GDF9*, 13B e 86B- *BMP15*) e Swiss White Alpine (35G- *GDF9*, 13B e 86B- *BMP15*).

A perspectiva futura é desenvolver e validar um painel de genotipagem de baixa densidade de SNPs relacionados ou potencialmente relacionados a taxa de ovulação/ número de borregos por parto na espécie ovina. Uma das prioridades, por exemplo, será selecionar mutações não conservativas nas regiões exônicas dos genes relacionados à taxa de ovulação (*GDF9*, *BMP15* e *BMPBR1-b*) identificadas no Capítulo 3. Adicionalmente os demais SNPs serão selecionados *in silico*, baseado em estudos já publicados (ex., Hanrahan et al., 2004; Silva et al., 2011; Demars et al., 2013; Souza et al., 2014).

Com essa ferramenta será possível, em um único teste, prever qual mutação causal dos três genes principais poderia estar associada a rebanhos com histórico de partos múltiplos. A presente equipe possui uma parceria com o Centro Nacional de Preservação e Recursos Genéticos (*National Center for Genetic Resources Preservation- NCGRP*), unidade do USDA-ARS, para realizar essa validação com amostras depositadas no banco de germoplasma americano.

Adicionalmente, o painel de baixa densidade em genes relacionados a taxa de ovulação desenvolvido a partir de SNPs identificados no trabalho atual será útil para expandir os estudos de interação entre genes/alelos envolvidos com características de prolificidade em diferentes raças do mundo. O painel permitirá a identificação de mutações e posterior seleção (a favor ou contra) de animais a partir de variantes associadas a efeitos de prolificidade em rebanhos de ovinos ao redor do mundo. A identificação desses polimorfismos ainda pode ser importante para qualificar germoplasma criopreservado em Bancos Genéticos Animais e agregar valor ao produto conservado.

REFERÊNCIAS BIBLIOGRÁFICAS

- DEMARS, J. et al. Genome-Wide Association Studies Identify Two Novel BMP15 Mutations Responsible for an Atypical Hyperprolificacy Phenotype in Sheep. **PLoS Genetics**, v. 9, n. 4, p. e1003482, 2013.
- HANRAHAN, J. P. Mutations in the Genes for Oocyte-Derived Growth Factors GDF9 and BMP15 Are Associated with Both Increased Ovulation Rate and Sterility in Cambridge and Belclare Sheep (*Ovis aries*). **Biology of Reproduction**, v. 70, n. 4, p. 900–909, 2004.
- LÔBO, R. N. B. et al. Economic values for production traits of Morada Nova meat sheep in a pasture based production system in semi-arid Brazil. **Small Ruminant Research**, v. 96, n. 2-3, p. 93–100, 2011.
- SILVA, B. D. M. et al. A new polymorphism in the Growth and Differentiation Factor 9 (GDF9) gene is associated with increased ovulation rate and prolificacy in homozygous sheep. **Animal Genetics**, v. 42, n. 1, p. 89–92, 2011.
- SOUZA, C. J. H. et al. Mutation in the protease cleavage site of GDF9 increases ovulation rate and litter size in heterozygous ewes and causes infertility in homozygous ewes. **Animal genetics**, v. 45, n. 5, p. 732–9, 2014.

ANEXO

CAPÍTULO 3

Tabela 1A- Haplótipos estimados para os genes *GDF9*, *BMP15*, *BMPRI-b* e sua distribuição por raças na Pop1.

Raças	Genes		Variantes na proteína.				Total de sequências (n)
	<i>GDF9</i> (HG)	<i>BMP15</i> (HB)	<i>BMPRI-B</i> (HBr)	<i>GDF9</i>	<i>BMP15</i>	<i>BMPRI-B</i>	
African White Dorper	2	3, 18, 86,87	1, 2, 3,4.		Leu10del L252P		4
Afshari	2, 5	6, 35, 36,37	5,6,7,8.			M64I	4
Awassi	5	6, 30	9,10.				2
Bangladeshi	2	8, 68, 84 e 85	11,12,13,14.		Leu10del L252P	G249R	4
Brazilian Creole	2, 7	3, 8, 9 e 10.	15,16,17,18.		Leu10del		4
Castellana	4, 13, 14, 42	23, 25, 40,41	19,20,21,22.	R87H			4
Changthangi	1, 2, 45	9, 11, 45, e 95	23,24.	E241K	P177A L252P		2
Cheviot	2, 10, 11, 30, 45	12, 13, 46 e 61	25,26,27,28.	Q80Stop E241K	P39R S346N		4
Churra	27, 28, 37, 38	70, 71, 76 e 77.	29,30,31,32,33,34.		P39R P177A S346N	M64I	6
Cine Capari	2	2 e 19	35,36.		Leu10del	M64I	2
Ethiopian Menz	2, 43	46 e 59	37,38.	R453C		M64I	2
Finnsheep	2, 4	3, 33, 49 e 90	39,40,41,42.		Leu10del		4
Garole	34	44 4 e 45	43,44.		L252P	G249R	2
Garut	2, 44	14 e 74	45,46,47,48.		Leu10del L252P	G249R	4
Gulf Coast native	2 e 7	26, 27, 63 e 64	49,50,51,52.				4
Indian Garole	7, 13	46 e 54	53,54.		L252P	G249R	2
Karakas	2, 14, 29	46 e 53	55,56,57,58.	R87H			4
Karya	2, 12	31 e 321	59,60.	R87H		V340A	2
Meat Lacaune	4, 33	3 e 39	61,62.		Leu10del		2
Merino	4, 7	50, 78, 79, 82 e 87	63,64,65,66,67,68.		Leu10del		6

CONTINUA...

Raças	Genes			Variantes na proteína.			Total de sequências (n)
	<i>GDF9</i> (HG)	<i>BMP15</i> (HB)	<i>BMPRI-B</i> (HBr)	<i>GDF9</i>	<i>BMP15</i>	<i>BMPRI-B</i>	
Milk Lacaune	2	80 e 81	69,70.				2
Morada Nova	1, 2, 7	4, 5, 6 e 7	71,72,73,74.	E241K	Leu10del	H335R	4
Namaqua Afrikaner	1, 3	17 e 18	75,76.	E241K	Leu10del L252P		2
Norduz	2, 7, 14	55, 61 e 62	77,78,79,80.	R87H			4
Ojalada	2, 7, 40	55, 72, 73 e 94	81,82,83,84.				4
Ovis canadensis	1, 14, 15, 16, 17, 22	3, 17, 50, 51, 56 e 57	85,86,87,88,89,90.	R87H, E241K	Leu10del S346N		6
Ovis dalli	2, 7, 18	6, 21 e 52.	91,92,93,94.			P171S	4
Poll Dorset	1, 25	52 e 58	95,96.	E241K			2
Romney	2, 4	9 e 25	97,98.				2
Ronderib Afrikaner	1, 7, 22	8, 42, 43 e 60	99,100,101,102.	E241K	Leu10del L252P		4
Sakiz	2	23, 24, 91 e 92	103,104,105,106.			M64I	4
Salz	7, 31	15, 16, 88, 89. E 93	107,108,109,110,111,112.	R87H	Leu10del		6
Santa Inês	2	1, 2 e 3	113,114,115,116.		Leu10del		4
Scottish Blackface	23, 24	46 e 50	117,118.	V332I			2
Sumatra	2, 39	48, 49, 53 e 54.	119,120,121,122.		Leu10del L252P	G249R	4
Swiss Mirror	2	23,28	123,124.				2
Swiss White Alpine	2, 3, 4, 7, 16, 21, 26 e 39.	3, 30, 38, 46, 47, 65, 66 e 83	125,126, 127, 128, 129, 130, 131, 132.		Leu10del P39R S346N		8
Texel	7, 36	46 e 67	133,134.				2
Tibetan, Eastern	2, 41	6 e 29	135,136.				2
Tibetan, Northern	2, 35	46 e 52	137,138.				2
Turkish Awassi	2, 7, 19, 20	6, 25, 53 e 69	139,140,141,142.	R87H			4
Valais Blacknose	2, 7	3 e 8	143,144.		Leu10del		2
Welsh Hardy Speckled Face	2	21 e 22	145,146.		L252P		2
Welsh Mountain, Dollgellau	2, 7	61 e 67	147,148.				2
Welsh mountain, Tregaron	8, 9	46 e 75	149,150.				2
Total	45	95	150	5	6	5	150