

Lílian Rosalina Gomes Silva

**Genômica comparativa para identificação de genes candidatos à resistência a endoparasitas gastrintestinais na espécie *Ovis aries***

**Teresina-PI**

**2019**

Lílian Rosalina Gomes Silva

**Genômica comparativa para identificação de genes  
candidatos à resistência a endoparasitas gastrintestinais  
na espécie *Ovis aries***

Dissertação apresentada ao Programa de Pós-Graduação em Ciência Animal, da Universidade Federal do Piauí, para obtenção do título de MESTRE EM CIÊNCIA ANIMAL. Área de concentração: Produção Animal.

Universidade Federal do Piauí

Orientador: José Lindenberg Rocha Sarmiento

Teresina-PI

2019

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

**S586g** Silva, Lílian Rosalina Gomes  
Genômica comparativa para identificação de genes candidatos  
à resistência a endoparasitas gastrintestinais na espécie *ovís aries*  
/ Lílian Rosalina Gomes Silva – 2019.  
96 f. : il.

Dissertação ( Mestrado) – Universidade Federal do Piauí, Pro-  
grama de Pós-Graduação em Ciência Animal, Teresina, 2019.  
Orientação: Prof. Dr. José Lindenberg Rocha Sarmiento

1. Ovinos 2. SNP 3. Alinhamentos genômicos 4. Parasitoses  
Helmínticas I. Título.

**CDD 636.3**

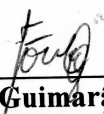
**GENÔMICA COMPARATIVA PARA IDENTIFICAÇÃO DE GENES  
CANDIDATOS À RESISTÊNCIA A ENDOPARASITAS GASTRINTESTINAIS  
NA ESPÉCIE *OVIS ARIES***

**LILIAN ROSALINA GOMES**

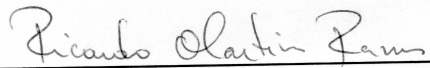
**Dissertação aprovada em: 26/02/2019**

**Banca Examinadora:**

  
\_\_\_\_\_  
**Prof. Dr. José Lindenberg Rocha Sarmiento (Presidente) / DZO/CCA/UFPI**

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

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

  
\_\_\_\_\_  
**Prof. Dr. Ricardo Martins Ramos (Externo) / IFPI**

# Lista de ilustrações

Figura 1 – Representação esquemática de um polimorfismo de nucleotídeo único(SNP <i>Single Nucleotide Polymorphism</i> ). A molécula de DNA do Indivíduo 1 difere da molécula de DNA do Indivíduo 2 em um par de bases individual (polimorfismo T/G). . . . .	17
Figura 2 – Representação esquemática de um alinhamento global e um alinhamento local. No alinhamento global as sequências de nucleotídeos são comparadas em toda a sua extensão. Já no alinhamento local, são localizados fragmentos de sequências que são mais similares. . . . .	25
Figura 3 – Representação da estatística de alinhamento feita pelo BLAST. O sistema de pontuação é regido pelas pontuações de <i>match</i> , <i>mismatch</i> e <i>gap</i> . O somatório destas pontuações é utilizado para calcular o <i>Score</i> de alinhamento. . . . .	28
Figura 4 – Saída do BLAST para o alinhamento entre duas sequências. Os <i>lengths</i> representam os tamanhos da <i>query</i> e <i>subject</i> . As pontuações de <i>match</i> , <i>mismatch</i> e <i>gaps</i> , irão determinar o <i>score</i> do alinhamento. A <i>Identities</i> é resultante da quantidade <i>matches</i> dentro do tamanho total do alinhamento. O <i>Expect</i> representa a confiabilidade do alinhamento. . . . .	29
Figura 5 – Localização do SNP no cromossomo e delimitação da janela para análise de genômica comparativa. . . . .	35
Figura 6 – Metodologia completa do trabalho: a) Etapa de seleção dos seguimentos a serem alinhados. b) Alinhamentos gerados pelo BLASTn. c) Análise dos genes oriundos dos melhores alinhamentos. . . . .	36
Figura 7 – Representação gráfica em escalas de pares de base (pb), de acordo com o filtro RAG=0,9, do alinhamento ocorrido entre genes das janelas e genes das espécies comparadas. . . . .	42
Figura 8 – Representação gráfica em escalas de pares de base (pb), de acordo com o filtro RAG=0,9, do alinhamento ocorrido entre genes das janelas e genes das espécies comparadas. . . . .	43

# Lista de tabelas

Tabela 1 – Bancos de dados mais comuns na área da bioinformática utilizados nesta pesquisa. . . . .	21
Tabela 2 – Programas BLASTs disponíveis para cada tipo de comparação de sequências. . . . .	26
Tabela 3 – Lista de SNPs com suas respectivas posições no cromossomo do <i>Ovis aries</i> e fenótipos de resistência ao qual foram associados. . . . .	35
Tabela 4 – Genes candidatos à resistência na espécie <i>O. aries</i> , e suas respectivas localizações em cada janela. . . . .	39
Tabela 5 – Quantidade de alinhamentos por janela antes e após os filtros de <i>E-value</i> , Identidade e RAG (Razão entre o tamanho do Alinhamento e tamanho do Gene). . . . .	40
Tabela 6 – Posições dos genes e dos segmentos das janelas que obtiveram similaridade identificada. . . . .	42
Tabela 7 – Genes encontrados na janela do SNP: OAR21_52583090.1 . . . . .	54
Tabela 8 – Genes encontrados na janela do SNP: OAR12_17005059.1 . . . . .	54
Tabela 9 – Genes encontrados na janela do SNP: OAR8_91253490_X.1. . . . .	54
Tabela 10 – Genes encontrados na janela do SNP: s23985.1. . . . .	55
Tabela 11 – Genes encontrados na janela do SNP: s26633.1. . . . .	55
Tabela 12 – Genes encontrados na janela do SNP: OAR2_148350187.1. . . . .	55
Tabela 13 – Genes com relação com processo imunológico-OAR21_52583090.1 . . . .	56
Tabela 14 – Genes sem relação com processo imunológico-OAR21_52583090.1 . . . .	57
Tabela 15 – Genes com relação com processo imunológico-OAR12_17005059.1 . . . .	58
Tabela 16 – Genes sem relação com processo imunológico-OAR12_17005059.1 . . . .	58
Tabela 17 – Genes sem relação com processo imunológico-OAR8_91253490_X.1 . . . .	58
Tabela 18 – Genes sem relação com processo imunológico- OAR2_148350187.1 . . . .	59
Tabela 19 – Genes sem relação com processo imunológico- OAR2_148350187.1 . . . .	60
Tabela 20 – Genes com relação com processo imunológico - s23985 . . . . .	60
Tabela 21 – Genes sem relação com processo imunológico - s23985 . . . . .	61

*Dedico este trabalho à minha mãe **Fracisca Girlene da Conceição**, ao meu irmão **José Milton Gomes** e ao meu amado esposo **Romuerre Rodrigues Veloso e Silva**.*

# Agradecimentos

À Deus, por ter me dado saúde, paciência e perseverança em todos os dias de minha vida, e por colocar pessoas especiais em meu caminho para me ajudarem nos momentos que mais precisei, e que com suas experiências me ensinaram a tomar boas decisões na escolha de qual caminho eu deveria trilhar.

À UFPI, por me proporcionar aperfeiçoar meus conhecimentos e poder contribuir com a pesquisa científica. Ao CNPq, pela concessão da bolsa para a realização desta pesquisa de estudo que foi de grande importância para a conclusão do curso. Ao Programa de Pós-Graduação em Ciência Animal, a todos os professores que me ajudaram a agregar grandes conhecimentos para que pudesse ocorrer o bom desenvolvimento da pesquisa.

Ao Prof. Dr. José Lindenberg Rocha Sarmiento pela orientação, dedicação, ensinamentos, compreensão e confiança, indispensáveis para realização deste estudo e para o meu crescimento profissional. Por sempre estar instigando o instinto de pesquisador em seus orientados.

Ao Prof. Dr. Romuere Rodrigues Veloso e Silva, por todo apoio, incentivo, paciência e dedicação desde a idealização do projeto, desenvolvimento e conclusão na área de bioinformática. Também pela grande parceria na vida, meu amado marido, sempre instigando meu desejo pelos estudos e pesquisas.

Ao Prof. Dr. Fábio Barros Britto, por toda disponibilidade e ajuda prestada nas orientações na parte genômica da pesquisa.

Ao Me. José Cleydson Ferreira da Silva, por toda a ajuda no desenvolvimento da pesquisa, trazendo seus conhecimentos na área de Bioinformática, e por toda disponibilidade, paciência e carinho.

A todo apoio dado por minha mãe Francisca Girlene da Conceição, e meu irmão José Milton Gomes, sempre me dando forças e muito amor para continuar perseverando nos estudos.

Aos meus amigos Bruna Lima, Jefson Moraes, Diniz, João Lopes, Geandro, Tâmara, Luciano, Laylson, Débora e Marcelo que estiveram sempre por perto incentivando e ajudando de toda forma possível para que eu conseguisse concluir o mestrado.

À TODOS O MEU MUITO OBRIGADA!!!



*O otimista é um tolo.  
O pessimista, um chato.  
Bom mesmo é ser um realista esperançoso.*  
**Ariano Suassuna**

# Resumo

**Resumo:** O controle de infecções helmínticas em ovinos tradicionalmente é feito com o uso de vermífugos, produto este que quando utilizado de forma indiscriminada pode levar a seleção de vermes resistentes. Neste sentido, é importante o desenvolvimento de formas alternativas de controle que minimizem o uso e gastos com anti-helmínticos na criação de ovinos. Como método de controle alternativo pode ser feita a identificação de ovinos que possuem genes de resistência a endoparasitas gastrintestinais para futura seleção genômica e melhoramento animal. Este trabalho teve por objetivo identificar no genoma da espécie *Ovis aries* genes e regiões candidatos à resistência a endoparasitas gastrintestinais por meio da análise e comparação genômica. Para identificação dos genes candidatos foram feitas buscas dentro de 6 janelas de 2Mb entorno de SNPs associados a fenótipos de resistência no genoma ovino, e posteriormente foram realizadas análises comparativas de DNA entre as janelas e genes das espécies *Capra hircus*, *Bos taurus*, *Sus scrofa*, *Mus musculus* e *Homo sapiens*. Para visualizar os genomas, verificar função de genes, comparar DNA e visualizar resultados, foram utilizados o Ensembl, UniProt, BLASTn e Kablammo, respectivamente. Os alinhamentos gerados pelo BLASTn passaram por filtros de identidade, RAG (Razão entre o tamanho do Alinhamento e tamanho do Gene) e *e-value*, onde foram identificados 281 alinhamentos significativos conforme esses filtros. A metodologia de investigação genômica norteadas pela informação de SNPs e da genômica comparativa possibilitaram a identificação dos genes FADD, LOC10111857, SLC22A18, ENSOARG0000000478, PRKN, CTSL, SLIT2, UBR3, KLHL41 e KLHL41, e de duas regiões localizadas respectivamente na posição 14658370-14659236 do cromossomo 12 e 46434090-46435478 do cromossomo 2. Ambos os genes e regiões passam a ser candidatos à resistência a endoparasitas gastrintestinais na espécie *Ovis aries*. As informações geradas nesta pesquisa servirão de base para que trabalhos mais aprofundados possam identificar os genes responsáveis pelo controle e expressão da resistência, que irão auxiliar na seleção e melhoramento de ovinos.

Palavras-chave: SNP. Alinhamentos genômicos. Parasitoses helmínticas. Ovinos.

# Abstract

**Abstract:** The traditional process to control helminth infections in sheep is using vermifuges, a product that can lead to the selection of resistant worms. In this sense, it is essential to develop alternative forms to control these infections. Thus, it is possible to identify individuals that have resistance genes for further genomic selection and animal breeding. This work aimed to identify genes and genomic regions that are candidates for resistance to gastrointestinal endoparasites in the genome of the *O. aries* species through analysis and comparative genomics. For this, we search six regions (windows) of 2Mb neighboring SNPs associated with resistance phenotypes in the ovine genome. Then, we performed a comparative DNA analyses between the windows and genes of the *C. hircus*, *B. taurus*, *S. scrofa*, *M. musculus* and *H. sapiens*. To acquire and visualize the genomes, verify gene function, compare DNA and visualize results, we used Ensembl, UniProt, BLASTn and Kablammo, respectively. We identified 281 significant alignments, generated by BLASTn passed by filters of identity filters, RAG (Ratio between Alignment size and Gene size) and *e-value*. Our proposed methodology enabled the identification of the genes: FADD, LOC10111857, SLC22A18, ENSOARG0000000478, PRKN, CTSL, SLIT2, UBR3, KLHL41 and KLHL41, and two regions located at chromosomes 12 and 2. Genes and regions become candidates for resistance to gastrointestinal endoparasites in the *O. aries* species.

**Key-words:** SNP. Genomic alignments. Helminthic parasites. Sheep.

# Sumário

<b>1</b>	<b>Introdução</b>	<b>12</b>
	<b>Introdução</b>	<b>12</b>
<b>2</b>	<b>Revisão de literatura</b>	<b>14</b>
2.1	A suscetibilidade de ovinos a endoparasitas	14
2.2	Marcador molecular do tipo SNP	16
2.3	Associação genômica ampla - <i>Genome-Wide Association Study</i> -GWAS	18
2.4	O papel da bioinformática na biotecnologia	20
2.4.1	O campo da genômica	22
2.5	Análise comparativa de sequências	22
2.5.1	Alinhamento de sequências biológicas	24
2.5.1.1	Tipos de alinhamentos	25
2.5.1.2	A ferramenta BLAST	26
2.5.1.3	O algoritmo do BLAST	27
2.5.1.4	A estatística do BLAST - como avaliar um alinhamento	27
2.5.1.5	A sensibilidade do BLAST	29
2.6	Trabalhos relacionados com esta pesquisa	30
<b>3</b>	<b>Genômica comparativa para identificação de genes candidatos à resistência a endoparasitas gastrintestinais na espécie <i>Ovis aries</i></b>	<b>31</b>
3.1	Introdução	33
3.2	Metodologia	34
3.2.1	Análise de janelas para a identificação de genes candidatos a resistência	35
3.2.2	Genômica comparativa para identificação de genes candidatos a resistência	35
3.2.3	Filtragem dos alinhamentos gerados pelo BLASTn <i>E-Value</i>	37
3.2.4	Filtragem por Identidade	38
3.2.5	Filtragem pela Razão entre o Tamanho do Alinhamento e o Tamanho do Gene (RAG)	38
3.3	Resultado e discussão	38
3.3.1	Genes candidatos à resistência identificados no genoma do <i>O. aries</i>	38
3.3.2	Alinhamentos ótimos do BLASTn	40
3.3.3	Identificação de genes nas janelas do <i>O. aries</i> que alinharam com genes das espécies comparadas	40
3.3.3.1	Identificação de genes candidatos à resistência via genômica comparativa	41
3.4	Conclusão	45
	<b>Referências</b>	<b>46</b>

3.5	Considerações finais . . . . .	52
-----	--------------------------------	----

<b>Apêndices</b>	<b>53</b>
------------------	-----------

<b>APÊNDICE A – Genes Encontrados nas Janelas do Ovis Aries . . .</b>	<b>54</b>
---	-----------

<b>APÊNDICE B – Alinhamentos que ocorreram entre genes . . . .</b>	<b>56</b>
--	-----------

# 1 Introdução

O Brasil tem uma longa tradição na criação de ovinos, com principal destaque para produção de carne, leite e lã. Segundo o [Instituto Brasileiro de Geografia e Estatística \(2018\)](#), no ano de 2017, o país apresentou um rebanho ovino de mais de 18 milhões de cabeças, com maior rebanho nas regiões Nordeste e Sul.

A produção de ovinos encontra-se em um quadro de expansão devido ao aumento do consumo de carne destes animais. No entanto, entraves na produção, como infecções parasitárias gastrointestinais, tem comprometido o aproveitamento econômico destes animais. Para tanto os programas de seleção e melhoramento genético de ovinos devem incluir além de características relacionadas à produção, a questão da resistência e resiliência a endoparasitas ([ALBERS et al., 1987](#); [AMARANTE et al., 2004](#)).

As infecções causadas por nematódeos gastrointestinais são uma das doenças parasitárias mais prevalentes que afetam pequenos ruminantes no mundo todo ([TORRES-ACOSTA; HOSTE, 2008](#)). Segundo [Melo et al. \(2003\)](#) e [Nova et al. \(2014\)](#) as helmintoses provocam grandes prejuízos, como, queda no desempenho dos animais, elevação dos custos de produção, e até mesmo a morte dos animais. Segundo [Waller et al. \(1996\)](#), o controle tradicional da doença, é feito com o uso de drogas anti-helmínticas, e o uso indiscriminado das mesmas podem levar a seleção de vermes resistentes, contaminação do meio ambiente e contaminação de produtos de origem animal. Além do que, a compra de vermífugos gera altos gastos para o produtor.

Uma forma alternativa para diminuir o uso de anti-helmínticos consiste em conhecer o material genético do animal de interesse, ou seja, identificar se o ovino possui genes de resistência a endoparasitas gastrointestinais. De acordo com [Daetwyler et al. \(2010\)](#), a identificação e estudo de marcadores moleculares tipo SNP (*Single Nucleotide polymorphism*) juntamente com estudos de Associação Genômica Ampla (GWAS - *Genome Wide Associations Study*), permitem identificar regiões no genoma associados com características de interesse. Os estudos de GWAS auxiliam no entendimento dos mecanismos genéticos que controlam características de herança complexa, como a resistência genética a verminose. A partir disso pode-se analisar se próximo a essas regiões de interesse há genes candidatos a característica.

De acordo com [BIAGIOTTI \(2016\)](#), os estudos de associação genômica tem por objetivo identificar variantes genéticas em um grande número de indivíduos, a fim de se descobrir associações estatísticas significativas com fenótipos de interesse. A associação de marcadores moleculares com características de interesse econômico, abrem caminhos para que se possa analisar regiões próximas ao marcador e identificar genes candidatos a

expressão e controle da característica. Porém, pode ocorrer da investigação revelar pouca ou nenhuma informação sobre genes próximos ao SNP. Para a resolução deste problema, pode-se lançar mão da genômica comparativa, que permite a comparação entre materiais genéticos de diferentes espécies e a extrapolação de informações entre as mesmas.

Os avanços ocorridos nas áreas da biologia molecular, genômica, estatística e ciências da computação, permitiram a criação da bioinformática, campo que alavancou e otimizou os estudos na área da genômica. Através da bioinformática criou-se ferramentas computacionais com grande poder de análise de dados volumosos advindos do sequenciamento genômico. A exemplo, pode-se destacar a ferramenta de busca e alinhamento local BLAST (*Basic Local Alignment Search Tool*). Através do BLAST é possível analisar um grande volume de dados e utiliza-los, por exemplo, na criação de mapas físicos e genéticos. Dessa forma, a comparação entre mapas possibilita a extrapolação de informações de uma espécie para outra através de alinhamentos genômicos. O alinhamento permite estabelecer uma correspondência entre duas sequências biológicas (HAGEN, 2001; DEROME et al., 2018).

A partir da problemática apresentada, objetivou-se com este trabalho, identificar genes candidatos à resistência a endoparasitas gastrintestinais em ovinos, analisando regiões específicas de DNA do *Ovis aries*, e comparando-as com os genes das espécies *Capra hircus*, *Bos taurus*, *Sus scrofa*, *Mus musculus* e *Homo sapiens*.

## 2 Revisão de literatura

### 2.1 A suscetibilidade de ovinos a endoparasitas

O setor da ovinocultura está em constante expansão no Brasil. Atualmente possui um efetivo de mais de 18 milhões de cabeças, com destaque para os rebanhos do Sul e Nordeste do país ([Instituto Brasileiro de Geografia e Estatística, 2018](#)). No entanto, a atividade pecuária sofre perdas econômicas devido a suscetibilidade dos animais a nematódeos gastrintestinais ([GIRÃO; MEDEIROS; GIRÃO, 1992](#)). As helmintoses causam prejuízos significativos na ovinocultura, restringindo o aumento da produção e causando perdas econômicas consideráveis ([WALLER, 1997](#)).

As estratégias de controle de helmintos no mundo todo baseiam-se quase que inteiramente no uso frequente de vermífugos (anti-helmínticos), que são cada vez mais considerados insustentáveis devido ao surgimento de múltiplos parasitas resistentes a drogas ([KAPLAN, 2004](#); [D., 1992](#)). É necessário o uso de métodos alternativos para fazer o controle do endoparasitismo. Acredita-se que a resistência anti-helmíntica é inevitável. Cada vez que um anti-helmíntico é administrado, o animal elimina parasitas suscetíveis e seleciona parasitas resistentes, que passam seus genes resistentes para a próxima geração de vermes ([SAYERS; SWEENEY, 2005](#)).

As doenças parasitárias, como as causadas pelo *Haemonchus contortus*, *Nematodirus*, *Fasciola hepatica*, *Dicrocoelium dendriticum*, *Eimeria* e *Amblyomma spp.*, entre outras, são importantes para a indústria ovina e são consideradas alguns dos maiores gargalos para o seu desenvolvimento ([MCMANUS et al., 2009](#)). Grandes perdas econômicas são causadas por essas doenças, como, retardo no desenvolvimento, perda de peso, redução do consumo de alimentos, diminuição da produção de leite, baixa fertilidade e, em casos de grandes infecções, altas taxas de mortalidade. O tratamento dessas doenças em ruminantes tem sido afetado pelo surgimento de cepas de nematoides resistentes aos anti-helmínticos.

Os ovinos podem ter várias espécies de nematelmintos os parasitando simultaneamente. A diversidade de parasitas nestes animais sofre influência de tratamentos anti-helmínticos, manejo e sanidade ambiental ([SANTIAGO; BENEVENGA; COSTA, 1976](#); [RAMOS et al., 2004](#)).

A principal espécie de parasita ovino no Brasil é o *Haemochus contortus*. As condições climáticas das regiões quentes e úmidas favorecem o parasitismo, demonstrando elevada patogenicidade e prolificidade. A fêmea deste verme libera até 5.000 ovos por dia, contaminando assim as pastagens. As populações desse nematódeo apresentam frequente resistência a vermífugos, e acometem principalmente os ovinos quem vivem em regiões com



clima tropical e subtropical. Este verme é hematófago e habita o abomaso dos ruminantes (CARVALHO et al., 2001; AMARANTE, 2014).

A resistência aos parasitas envolve mecanismos de resposta passivos e ativos. Barreiras físicas e químicas constituem os mecanismos passivos, que impedem a entrada ou alojamento do parasita no corpo do hospedeiro. Já a resistência ativa compreende a resposta imunológica ativa e/ou adaptativa, desencadeada em resposta as infecções (SADDIQI et al., 2011). Neste caso, os antígenos produzidos pelos parasitas ativam o sistema imunológico, contando com a ajuda de células de defesa especializadas que produzem inúmeras substâncias, como citocinas e imunoglobulinas (anticorpos) (AMARANTE, 2014).

Os nematódeos induzem resposta mediada por linfócitos Th2, dependente de anticorpos, enquanto que parasitas intracelular, como o caso dos vírus, induzem resposta Th1. Estes dois tipos de respostas são importantes, pois os animais são expostos a diferentes patógenos. O hospedeiro resistente será o que melhor elaborar uma resposta imunológica eficiente, ou seja, elaborar resposta protetora tipo Th2. As células do tipo Th2 produzem citocininas que estimulam a proliferação de células inflamatórias, como eosinófilos e mastócitos. As citocininas também estimulam a produção de imunoglobulinas contra os parasitas, tais como a IgE e IgA. Há um aumento na produção de muco, que também podem ter ação antiparasitária (ANTHONY et al., 2007).

Os ruminantes podem indicar imunidade contra nematódeos adultos através da expulsão da população adulta destes vermes. Outros sinais incluem alterações da morfologia dos parasitas e redução da fecundidade das fêmeas. No caso da resistência contra as larvas, acontecerá a expulsão ou inibição do desenvolvimento dos parasitas imaturos (hipobiose) (BALIC; BOWLES; MEEUSEN, 2000).

Os animais que são acometidos por processos infecciosos possuem menor desempenho quando comparados com animais que estão livres deste estágio. A habilidade do animal suportar a doença sofre influência de fatores ambientais, onde o estado nutricional e fisiológico do animal terão grande peso para que a resposta imunológica contra patógenos funcione adequadamente (BISHOP; STEAR, 2003). Apesar das infecções poderem resultar em morte do hospedeiro, a principal consequência dessas verminoses tem sido a redução da produtividade do rebanho (AMARANTE, 2014).

A seleção de animais resistentes a determinada espécie de nematelminto pode contribuir para que o mesmo animal também seja resistente a uma outra espécie. A exemplo, pode-se citar a raça Merino que foi selecionada para resistência contra *H. contortus* e posteriormente também demonstrou resistência contra *T. colubriformis* (SRETER; KASSAI; TAKACS, 1994).

A aquisição e expressão de imunidade contra nematóides gastrintestinais é geneticamente controlada e varia entre e dentro de raças (BISHOP; WOOLLIAMS, 2004;

STEAR; MURRAY, 1994). A proporção de animais considerados resistentes ou suscetíveis a nematoides é influenciada pela idade e raça (AMARANTE et al., 2004). A variação genética na resistência a parasitas internos foi demonstrada em numerosas espécies, incluindo humanos e animais (KLOOSTERMAN; PARMENTIER; PLOEGER, 1992; QUINNELL, 2003; BISHOP; MORRIS, 2007). No entanto, a arquitetura genética subjacente a tais características é pouco compreendida (KEMPER et al., 2011). Os aspectos genéticos relacionados ao controle de helmintos dependem de diversos fatores, como, a raça, a influência do ambiente na expressão dos genes, a probabilidade da característica ser transmitida as próximas gerações, interação com outras características de interesse econômico, assim como a interação com marcadores genéticos, incluindo lócus de características quantitativas (QTLs) e genes de efeito maior.

A implementação de programas de melhoramento genético torna-se um dos fatores mais importantes a serem levados em consideração, já que a interação parasita-hospedeiro ocorre em vários níveis. Os esquemas de seleção podem conferir resistência ou tolerância à infecção. O termo resistência refere-se à capacidade do hospedeiro resistir à infecção, enquanto que, com tolerância, o hospedeiro é infectado pelo patógeno, mas sofre pouco efeito adverso. Por exemplo, quando o objetivo é prevenir a disseminação da doença para outras populações (como no caso de doenças zoonóticas), é necessária resistência à doença em vez de tolerância.

A transmissão da infecção geralmente é afetada quando a melhora genética é feita na resistência do hospedeiro à infecção. Embora a melhora genética da tolerância possa reduzir os sinais clínicos da doença, ela pode não reduzir a transmissão da infecção para outros animais (GIBSON; BISHOP, 2005).

A utilização de ovinos resistentes a verminoses tem grande impacto do que diz respeito à epidemiologia das infecções por nematódeos gastrintestinais. Mesmo as raças que apresentam elevada produtividade e que ainda sim são consideradas suscetíveis, possuem variabilidade genética em relação a resposta imunológica contra os patógenos, e isso é suficiente para se entender que a resistência a verminoses deve ser incluída como um dos critérios de seleção animal nos programas de melhoramento genético (AMARANTE, 2014).

## 2.2 Marcador molecular do tipo SNP

Os marcadores moleculares são peças-chave na genética, e surgiram devido à necessidade de se detectar polimorfismos diretamente no DNA (*Deoxyribonucleic Acid*) (VIGNAL et al., 2002). Segundo Al-Samarai e Al-Kazaz (2015), um marcador genético é um gene ou sequência de DNA com uma localização conhecida num cromossomo e associada a um gene ou característica particular. É descrito como uma variação, que pode surgir devido a mutação ou alteração nos lócus genômicos que são observados. Um marcador

genético pode ser uma sequência curta de DNA, tal como uma sequência envolvendo uma única troca de pares de bases (polimorfismo de nucleotídeo único, SNP), ou uma longa, como mini e micro satélites.

Os últimos anos testemunharam um grande interesse pelos marcadores moleculares, revelando polimorfismo ao nível do DNA, pois eles desempenham um papel importante nos estudos de genética animal (AL-SAMARAI; AL-KAZAZ, 2015).

Em 1996, foi proposta uma nova tecnologia de marcadores moleculares chamada SNP (*Single-nucleotide polymorphism*). De acordo com Yang et al. (2013), este marcador consiste em um polimorfismo de sequência causado por uma única mutação de nucleotídeo em um locus específico na sequência de DNA (Figura 1).

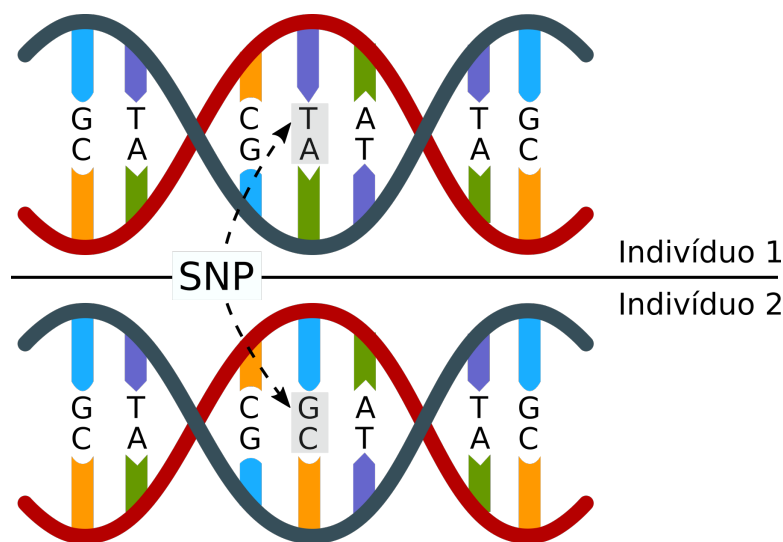


Figura 1: Representação esquemática de um polimorfismo de nucleotídeo único (SNP *Single Nucleotide Polymorphism*). A molécula de DNA do Indivíduo 1 difere da molécula de DNA do Indivíduo 2 em um par de bases individual (polimorfismo T/G).

Segundo Brookes (1999), os SNPs diferenciam-se dos demais tipos polimorfismos por serem nucleotídeos conhecidos e com posições definidas, onde geralmente são detectados através de comparações de sequências de DNA. Os SNPs são classificados de acordo com o tipo de variação de nucleotídeo em transições, purina-purina (A/G) ou pirimidina-pirimidina (C/T), e transversões, purina-pirimidina ou pirimidina-purina (A/C, A/T, G/C, G/T).

Atualmente, os marcadores SNP são uma das abordagens mais populares. Esta popularidade baseia-se em algumas propriedades preferenciais, pois são abundantes no genoma, geneticamente estáveis e passíveis de análise automatizada de alto rendimento. O conceito de SNP surgiu basicamente da necessidade de densidades muito altas de marcadores genéticos para os estudos de doenças multifatoriais (VIGNAL et al., 2002).

Diversos estudos confirmatórios apontam que os SNPs podem ser usados para

identificar genes que estão associados a doenças severas, resistência a drogas, e na identificação de fenótipos de interesse por genética de associação/desequilíbrio de ligação (SACHIDANANDAM et al., 2001; BADER, 2001).

Os SNPs com informação podem ser usados na seleção assistida (PAVY et al., 2006) e incorporados em programas de melhoramento de plantas e animais (DU et al., 2009). Além disso, os SNPs também podem ser úteis na construção de mapas genéticos de alta resolução, diagnóstico genético, análises filogenéticas (RAFALSKI, 2002) ou em estudos de história e genética de populações (STEPHENS, 2001).

Importantes avanços foram alcançados na genotipagem animal utilizando chips de SNPs com plataformas já disponíveis para a detecção simultânea de milhares de SNPs (COUTINHO; ROSÁRIO; JORGE, 2010). Esses chips de SNPs estão promovendo uma revolução na genômica animal, permitindo a varredura do genoma de uma população experimental para milhares de SNPs simultaneamente, a custos relativamente mais baixos e menor tempo em comparação com o que é feito no uso de microssatélites. Assim, permitindo fazer a associação genômica ampla, além de possibilitar a seleção genômica.

A dificuldade encontrada na utilização ampla de marcadores moleculares em programas de melhoramento genético está na identificação dos genes que controlam as características de interesse econômico. Apesar do sequenciamento completo do genomas de muitas espécies, a identificação de milhares de genes que conferem resistência a doenças, a ecto e endoparasitas, ao calor, a melhor conversão alimentar e maior eficiência reprodutiva, entre outras características, ainda são desconhecidos. A limitação não é mais quanto ao conhecimento da sequência de genes presentes nos animais domésticos, mas sim identificar os genes responsáveis pelo controle das características (COUTINHO et al., 2007).

## 2.3 Associação genômica ampla - *Genome-Wide Association Study*-GWAS

Marcadores moleculares têm sido uma ferramenta poderosa para identificar mutações que influenciam características controladas por um gene ou alguns genes. No entanto, a maioria das características de interesse econômico apresentam um padrão de herança poligênica, sendo determinadas por numerosos genes de grande e/ou pequeno efeito individual e sob forte influência de fatores ambientais. A identificação de alelos associados a essas características complexas nas populações só foi possível após o desenvolvimento e localização de marcadores moleculares polimórficos no genoma, particularmente microssatélites e polimorfismos de nucleotídeo único (SNP), permitindo a construção de mapas genéticos saturados e, conseqüentemente, mapeamento de lócus de caracteres quantitativos (*Quantitative trait Locus -QTL*)(COUTINHO; ROSÁRIO; JORGE, 2010).

O estudo de associação genômica é uma técnica para a identificação de genes causais para características importantes na pecuária. O GWAS usa variações de sequência (principalmente polimorfismos de nucleotídeo único, SNPs) em todo o genoma, juntamente com o fenótipo e informações de pedigree, para realizar análises de associação e para identificar genes ou elementos reguladores que são importantes para as características de interesse. O GWAS tornou-se viável tanto em humanos quanto em animais domésticos, como resultado do desenvolvimento de grandes coleções de SNPs e do desenvolvimento de métodos eficazes em termos de custo para a análise de SNPs em larga escala (HIRSCHHORN; DALY, 2005). O GWAS é uma técnica ideal para descobrir os principais genes para características complexas e é uma nova maneira de estudar o mecanismo genético destas características.

Estudos de associação genômica ampla são uma ferramenta importante para identificar loci de suscetibilidade a doenças e têm sido bem sucedidos em detectar a associação de vários SNPs com doenças complexas (SAXENA et al., 2007; PLENGE et al., 2007). A extração da percepção biológica de GWAS e a compreensão dos princípios subjacentes aos fenômenos complexos que ocorrem em vários caminhos biológicos continuam sendo um grande desafio. A abordagem comum do GWAS é selecionar dezenas dos SNPs mais significativos de uma lista para futuras investigações (LESNICK et al., 2007).

O GWAS foi usado pela primeira vez na análise de doenças humanas e grandes progressos foram feitos. O GWAS foi estendido para o campo da genética e reprodução de animais domésticos quando as sequências genômicas estavam disponíveis para várias espécies domésticas e descobriu-se formas de acessar um grande número de SNPs, sendo estes um subproduto de sequenciamento ou em subsequente re-sequenciamento. Existem muitos tipos de chips SNP comerciais disponíveis para bovinos (50.000 SNPs; Illumina BovineSNP50 BeadChip), cães (22.362 SNPs; Illumina CanineSNP20 BeadChip), ovinos (56.000 SNPs; Illumina OvineSNP50 BeadChip), suínos (60.000 SNPs; Illumina PorcineSNP60 BeadChip), cavalos (54.602 SNPs). ; Illumina EquineSNP50 BeadChip) e galinhas (60.000 SNPs; Illumina ChickenSNP60 BeadChip). Embora a aplicação de GWAS a animais domésticos tenha ocorrido apenas recentemente, tem havido uma série de resultados relatados (ZHANG et al., 2012).

Uma suposição feita na análise de GWAS é que associações significativas podem ser detectadas porque os SNPs estão em desequilíbrio de ligação (LD) com as mutações causativas para as características de interesse. A alta densidade de marcadores SNP no chip usado em GWAS foi suficiente para identificar o LD entre marcadores SNP e mutações causativas. Durante os últimos anos, vários exemplos de GWAS bem sucedidos em animais domésticos, incluindo bovinos, suínos, cavalos, cães, ovelhas e galinhas, foram relatados (ZHANG et al., 2012).

O primeiro relato do uso de GWAS em ovinos foi feito nos tipos de chifre de

JOHNSTON *et al.* (2011). Um estudo de associação genômica ampla foi conduzido utilizando 36.000 SNPs e determinou que o principal candidato genético para chifres seria RXFP2, um gene autossômico com envolvimento conhecido na determinação de características sexuais primárias em humanos e camundongos (FERLIN *et al.*, 2008; YUAN *et al.*, 2010; FENG *et al.*, 2009). Evidências de SNPs adicionais em torno de RXFP2 suportam um novo modelo de herança do tipo chifre em ovelhas Soay, e pela primeira vez ovelhas com o mesmo fenótipo de chifre, mas diferentes genótipos subjacentes podem ser identificados. Além disso, o RXFP2 demonstrou ser um locus de característica quantitativa aditivo (QTL) para o tamanho do chifre em machos com chifres normais, representando até 76% da variação genética aditiva nessa característica. Esta descoberta contrasta marcadamente com os GWAS de características quantitativas em humanos e algumas espécies modelo, onde é frequentemente observado que loci mapeados apenas explicam uma proporção modesta da variação genética global.

O outro estudo de GWAS em ovinos foi relatado por Zhao *et al.* (2011), que no mesmo ano usaram o mesmo Illumina OvineSNP50 BeadChip como JOHNSTON *et al.* (2011). Este estudo foi focado na herança de raquitismo em ovelhas Corriedale. Um GWAS foi realizado em 20 ovelhas relacionadas, compreendendo 17 indivíduos afetados e 3 portadores. Uma região homozigótica que incluiu 125 loci SNP consecutivos foi identificada em todas as 17 ovelhas afetadas, cobrindo uma região de 6 Mb no cromossomo 6 ovino. Havia 35 genes nessa região; o gene da proteína 1 da matriz dentinária (DMP1) foi sequenciada e uma mutação sem sentido, 250C / T, foi identificada no exon 6. Esta mutação introduziu um códon de parada (R145X) e pode truncar os aminoácidos C-terminais. A genotipagem por PCR-RFLP para esta mutação mostrou que todas as 17 ovelhas afetadas tinham o genótipo “TT”; os 3 portadores foram “CT”; 24 ovelhas relacionadas fenotipicamente normais eram “CT” ou “CC”; 46 ovelhas controle não relacionadas normais de outras raças eram todas “CC”. Os outros SNPs em DMP1 não eram concordantes com o raquitismo hereditário e todos podem ser descartados como candidatos. Pesquisas anteriores mostraram que mutações no gene DMP1 são responsáveis por raquitismo hipofosfatêmico autossômico recessivo em humanos (KOSHIDA *et al.*, 2010). Camundongos knockout *Dmp1* exibem fenótipos de raquitismo (FENG *et al.*, 2006). Portanto, acredita-se que a mutação R145X em DMP1 seja responsável por raquitismo hereditário em ovelhas Corriedale.

## 2.4 O papel da bioinformática na biotecnologia

Inúmeros organismos já tiveram suas sequências genéticas determinadas, e esses estudos produzem imensas quantidades de dados de sequências. Catalogar, armazenar, recuperar e analisar esse imenso conjunto de dados são os principais desafios da genética moderna. A bioinformática é um campo emergente que consiste em biologia molecular e



ciência computadorizada centrada no desenvolvimento de bancos de dados, pesquisa de algoritmos no computador, programas de previsão de genes e outras ferramentas analíticas que são usadas para dar sentido aos dados de DNA, RNA e sequência de proteínas. A bioinformática desenvolve e aplica essas ferramentas para "garimpar os dados", extraíndo as informações úteis dos projetos de sequenciamento. O desenvolvimento e uso de algoritmos e programas de computação para analisar dados de DNA e sequências de proteínas ajudaram a tornar a biologia molecular um campo mais quantitativo. Os dados de sequências em bancos de dados públicos, gratuitamente pesquisáveis com uma ligação com a Internet, permitiram que cientistas e estudantes de todo o mundo tenham acesso a esse extraordinário recurso (PIERCE, 2011).

Vários bancos de dados foram estabelecidos para a coleta e análise de informações das sequências de DNA e proteínas. Os bancos de dados primariamente contêm informações sobre a sequência, juntamente com informações que descrevem a fonte da sequência e sua determinação. Os bancos de dados secundariamente contêm os resultados de análises feitas nos dados primários de sequências, tais como informações sobre padrões particulares de sequências, variações, mutações e relações evolutivas. Alguns bancos de dados de bioinformática amplamente usados são listados na Tabela 1.

Tabela 1: Bancos de dados mais comuns na área da bioinformática utilizados nesta pesquisa.

Nome	Descrição	URL
GenBank	Informação primária da sequência de DNA mantida pelo U.S National Institutes of Health	<a href="http://www.ncbi.nlm.nih.gov/Genbank/">http://www.ncbi.nlm.nih.gov/Genbank/</a>
EMBL-Bank	Informação primária da sequência de DNA mantida por pesquisadores europeus	<a href="http://www.ebi.ac.uk/embl/">http://www.ebi.ac.uk/embl/</a>
UniProt	Dados de sequências de proteínas e outras informações sobre proteínas de uma variedade de organismos	<a href="http://www.ebi.ac.uk/uniprot/">http://www.ebi.ac.uk/uniprot/</a>

Após o sequenciamento de um genoma é feita a identificação dos genes potenciais dentro da sequência. Um gene identificado deve ser anotado, ou seja, ligar sua sequência a outra informação sobre sua função e expressão, proteína que ele codifica e a informação em genes similares em outras espécies (uso da genômica funcional). Estão disponíveis programas de computação para determinar se sequências similares já foram encontradas, seja na mesma espécie ou em espécies diferentes. O mais amplamente usado desses programas é o *Basic Alignment Search Tool (BLAST)*.

Bancos de dados adicionais contêm informação sobre a diversidade da sequência, como e onde um genoma varia entre organismos individuais. Os genomas são mapeados

em inúmeros SNPs, e a informação sobre a sequência dessas variantes em populações diferentes são coletadas. As informações importantes sobre os padrões de expressão de milhares de genes encontrados nos genomas também são compiladas (PIERCE, 2011).

### 2.4.1 O campo da genômica

O estudo científico dos processos biológicos pelo ponto de vista de todo genoma, é conhecido como genômica, e originou-se no Projeto Genoma Humano (HGP). Esse projeto, iniciado em 1980, foi audacioso, pois na época não existiam nem as tecnologias para gerar grandes quantidades de sequências de DNA, nem o poder computacional para analisar esse grande volume de dados (BOWMAN; SANDERS, 2014).

A genômica é dividida em: genômica estrutural que se concentra no sequenciamento dos genomas inteiros e na catalogação, ou anotação, das sequências dentro de um determinado genoma; a genômica evolutiva que é a comparação dos genomas, tanto dentro quanto entre espécies; e a genômica funcional, que utiliza as sequências genômicas para compreender a função gênica em um organismo. Juntas, essas três abordagens contribuem para o objetivo final de compreender o papel de cada gene contido em um determinado genoma (BOWMAN; SANDERS, 2014).

No setor de produção animal, por exemplo, a genômica permite a produção de genótipos superiores, fazendo-se o uso de seleção assistida por marcadores, transgenia e etc. Isso permite a produção de animais resistentes a doenças, mais adaptados às necessidades ambientais, com carnes melhores, com menos gordura, clones de animais com padrões excepcionalmente bons, seja na produção de carnes ou leite e até mesmo animais que funcionem como biofábricas de fármacos ou outros produtos. Toda informação advinda de estudos genômicos são essenciais para incorporação dos animais em programas que visam melhorar a produção e qualidade do produto final.

Nos pequenos ruminantes (ovinos e caprinos), as pesquisas genômicas estão trazendo novas perspectivas para o setor produtivo, possibilitando a aplicação dos resultados em ações para o melhoramento genético animal. As técnicas facilitam a identificação dos melhores animais de acordo com as características de interesse econômico, como a qualidade de leite, carne, lã, resistência a verminoses e preservação de padrões típicos de raças.

## 2.5 Análise comparativa de sequências

Atualmente há um grande volume de dados genéticos advindos do crescente número de organismos sequenciados. Sem dúvida, cria-se muita expectativa ao falar sobre o sequenciamento completo de genoma. Porém, argumenta-se que o genoma isolado de um único organismo não revela muito por si só.



Os princípios fundamentais da genômica comparativa são diretos. Características comuns de dois organismos serão muitas vezes codificadas dentro do DNA que é conservado entre as espécies. Mais precisamente, as sequências de DNA que codificam as proteínas e RNAs responsáveis por funções que foram conservadas a partir do último ancestral comum devem ser preservadas nas sequências genômicas contemporâneas. Da mesma forma, as sequências de DNA que controlam a expressão de genes que são regulados de forma semelhante em duas espécies relacionadas também devem ser conservadas. Por outro lado, sequências que codificam (ou controlam a expressão) proteínas e RNAs responsáveis por diferenças entre espécies serão elas próprias divergentes (HARDISON, 2003).

Genomas e genes precisam ser estudados em comparação com outras espécies (ou subespécies, ou estirpes), no contexto filogenético do processo evolutivo (CLARK, 1998). A genômica comparativa é uma abordagem holística de larga escala que compara dois ou mais genomas para descobrir as semelhanças e diferenças e estudar a biologia dos genomas individuais. A genômica comparativa aplica-se a genomas completos ou regiões sintênicas de diferentes espécies, subespécies diferentes ou estirpes diferentes da mesma espécie. A pesquisa comparativa em genômica inclui o desenvolvimento de ferramentas computacionais e o uso de ferramentas para analisar genomas para descobertas biológicas (WEI et al., 2002).

Nos últimos cinco anos, houve uma explosão de avanços computacionais e biológicos no campo jovem da genômica comparativa. As aplicações práticas da genômica comparativa são muitas e seu impacto científico é profundo.

No estudo genômico, é de suma importância entender a diferença entre os termos homologia e similaridade. Homologia implica ancestralidade comum de dois genes ou produtos gênicos. Semelhança é o que podemos medir a partir do alinhamento de sequências ou estruturas. Similaridade pode ser usada como evidência para homologia, mas não implica necessariamente homologia (WEI et al., 2002). A análise comparativa do genoma é frequentemente realizada entre espécies modelo e não-modelo (PATERSON et al., 2000; SCHMIDT, 2002; HALL; FIEBIG; PREUSS, 2002).

A essência da análise comparativa do genoma é a extrapolação da informação de um organismo para outro. O mapeamento comparativo e a análise comparativa de sequências são os componentes-chave da genômica comparativa. Mapeamento comparativo estabelece as relações sintênicas entre genomas de espécies diferentes, auxiliando na consolidação do mapa genético, a verificação de loci de características quantitativas (QTL), a identificação de genes candidatos subjacente QTL, e uma melhor compreensão da evolução do genoma (SANKOFF; NADEAU, 2000; KLIEBENSTEIN; GERSHENZON; MITCHELL-OLDS, 2001; MURPHY; STANYON; O'BRIEN, 2001; ZHANG et al., 2001).

Diferentes estratégias podem ser usadas para detectar genes e sequências de DNA e comparar suas posições e arranjos em diferentes espécies. Classicamente, os genes

eram reconhecidos pelos seus efeitos sobre o fenótipo, e um mapa genético foi construído observando-se os padrões de segregação entre os descendentes dos pais, diferindo em duas ou mais características. Os mapas genéticos ainda dependem de diferenças herdadas, mas estes podem ser reconhecidos tanto no nível molecular como no fenotípico. Mapas físicos descrevem a localização física de um gene ou uma sequência de DNA em uma região cromossômica, com maior ou menor resolução. Mapas em escala molecular podem ser construídos com a sequência base da região em seu ponto final (GRAVES, 1998).

De acordo com Graves (1998), diferentes tipos de mapas requerem diferentes tipos de marcadores genéticos. O mapeamento genético (*linkage*) se baseia na variação intraespecífica (alelos) em um locus gênico específico, enquanto o mapeamento genético de células somáticas requer variação interespecífica, e técnicas de mapeamento físico requerem uma sonda de DNA clonada. A disponibilidade de diferenças moleculares e fenotípicas entre os alelos aumentou enormemente a facilidade do mapeamento genético, especialmente em humanos, onde a informação é compilada a partir de cruzamentos limitados e não controlados. É importante ressaltar que essa disponibilidade também facilitou o alinhamento de mapas genéticos e físicos da mesma espécie e a movimentação entre eles no isolamento físico de um gene alvo.

Pode-se perceber que ao passo que ocorrem as comparações das sequências genômicas, é fornecida uma visão altamente detalhada de como os organismos estão relacionados entre si a nível genético. Para os pesquisadores que fazem genômica comparativa, os animais domésticos possuem uma grande vantagem sobre as demais espécies, pois os mesmos possuem uma longa história de estudo. Os pesquisadores conhecem o histórico de desenvolvimento, resistência a doenças, e conhecem diferentes formas de trabalhar experimentalmente os animais domésticos.

### 2.5.1 Alinhamento de sequências biológicas

A busca por similaridade de sequências é uma das atividades mais importantes da bioinformática, e frequentemente fornece a primeira evidência para a função de um novo gene sequenciado ou pedaço de sequência. Para conseguir esta similaridade é preciso fazer o alinhamento das sequências biológicas.

Após o sequenciamento, é preciso identificar e agregar informação a sequência. A melhor maneira de fazer isto, é comparar a sequência com sequências já conhecidas. Para isso, é feito o alinhamento par a par com um banco de dados. O que realmente importa é o grau real de similaridade, ou seja, alinhamentos estatisticamente significantes conforme a **identidade** (número que indica a quantidade de nucleotídeos alinhados); **similaridade** (considera a probabilidade do alinhamento ter ocorrido por acaso (*e-value*) e considera todos os outros possíveis alinhamentos); e **homologia** (refere-se ao fato de dois genes partilharem ou não uma história evolutiva comum).

No contexto de evolução as sequências de DNA sofrem mutações, e estas modificações locais entre os nucleotídeos podem ser: Inserções: inserção de uma base ou várias bases na sequência; Deleções: deleção de uma base ou mais bases na sequência; e Substituições: substituição de uma base por outra. Portanto, um programa de alinhamento de sequências biológicas tem que levar em consideração essas mutações.

O conceito básico para a seleção de uma boa sequência de alinhamento é simples. As duas sequências são combinadas aleatoriamente. A qualidade da pontuação é avaliada e pontuada. Em seguida uma sequência é movida em relação a outra e a combinação é pontuada novamente, até que seja obtida a melhor pontuação de alinhamento.

### 2.5.1.1 Tipos de alinhamentos

Os alinhamentos podem ser global ou local, a escolha é feita de acordo com o objetivo de análise. O alinhamento global compara as sequências em sua extensão total. É apropriado para comparação de sequências que se esperam conter similaridade em toda a sua extensão. O alinhamento maximiza as regiões de similaridade e minimiza os *gaps*. A diferença de alinhamento local e global podem ser observadas na Figura 2.

Na Figura 2, ambos os alinhamentos correspondem ao pareamentos de duas sequências onde uma letra em uma sequência (correspondente de um nucleotídeo ou aminoácido) é pareado com outra letra ou *gap* (-) da outra sequência. Cada alinhamento possui um valor de pontuação (*score*) que indica o quanto o alinhamento tem posições similares/idênticas. Esse *score* é calculado por algoritmos que computam o valor do pareamento de cada letra no alinhamento e depois somam esses valores. O cálculo do *score* de cada alinhamento é feito com base no tipo de sequência, proteína ou DNA.

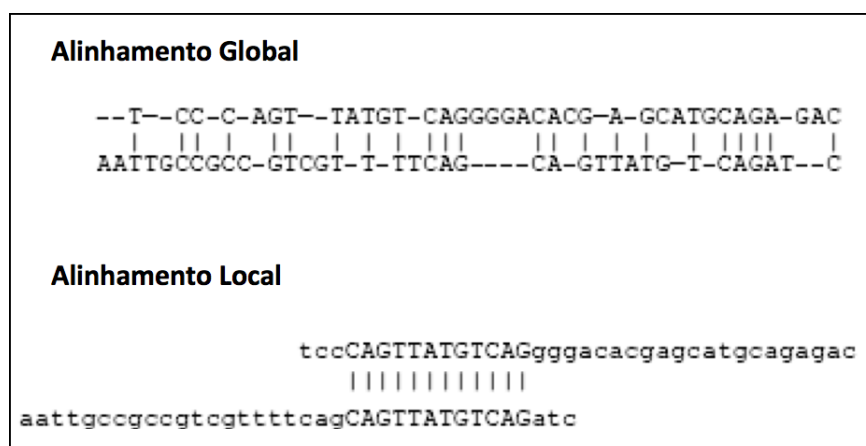


Figura 2: Representação esquemática de um alinhamento global e um alinhamento local. No alinhamento global as sequências de nucleotídeos são comparadas em toda a sua extensão. Já no alinhamento local, são localizados fragmentos de sequências que são mais similares.

### 2.5.1.2 A ferramenta BLAST

A Ferramenta de Pesquisa de Alinhamento Local Básica (BLAST) é provavelmente a ferramenta de pesquisa de similaridade mais popular. O Centro Nacional de Informações sobre Biotecnologia (NCBI) introduziu o BLAST pela primeira vez em 1989. O NCBI continuou a manter e atualizar o BLAST desde a primeira versão. O BLAST realiza comparações, concentrando-se em alinhamentos locais. Essa abordagem tem sido muito bem-sucedida, pois frequentemente apenas regiões localizadas, como sites ativos, são preservadas entre sequências (NCBI, 2008).

O BLAST pode ser usado de forma local, ou seja, instalado no desktop, no qual seu uso é mais vantajoso que a sua versão online, pois pode ser usado bancos de dados próprios e atualizações. As facilidades são a velocidade de buscas, automatização, dados locais e independente de internet.

As informações sobre as sequências, tanto DNA quanto proteínas, são armazenadas no formato de apenas uma letra para cada nucleotídeos (A, T, C e G) ou aminoácidos. As informações sobre cada sequência *query* e *subject* no banco de dados são armazenadas no formato FASTA. O BLAST possui um *default* de palavras, (ex.: DNA= 11 nt) onde os alinhamentos tem que ser exatos, e tem como saída o *score*, *e-value*, *identity*, *match*, *mismatch* e *gap*.

Cinco programas diferentes estão disponíveis: blastn, blastp, blastx, tblastn e tblastx. Essencialmente, todas as combinações de sequências de consulta de DNA ou proteína com buscas contra bases de dados de DNA ou proteínas são possíveis. A Tabela 2 traz a descrição destes programas.

Tabela 2: Programas BLASTs disponíveis para cada tipo de comparação de sequências.

Programa	Nome da Tarefa	Descrição
blastp	blastp	BLASTP tradicional para comparar uma consulta de proteína com um banco de dados de proteínas
	blastp-short	BLASTP otimizado para consultas com menos de 30 resíduos
blastn	blastn	BLASTN tradicional que exige uma correspondência exata de 11
	blastn-short	Programa BLASTN otimizado para sequências menores que 50 bases
	megablast	Megablast tradicional usado para encontrar sequências muito semelhantes (por exemplo, intraespécies ou espécies intimamente relacionadas)
	dcmegablast	Megablast descontínuo usado para encontrar sequências mais distantes (por exemplo, interespecies)

Na Tabela 2, encontram-se os tipos de BLAST disponíveis, que devem ser escolhidos de acordo com a necessidade da pesquisa. Cada BLAST possui um tipo de tarefa, a mesma é escolhida a partir do tamanho dos resíduos, tipo de material a ser comparado e finalidade. A descrição da funcionalidade de cada BLAST e tarefa desempenhada são indispensável para se saber em qual categoria o trabalho se encaixa, e só assim poder executar o alinhamento comparativo de forma confiável e condizente com os objetivos que almejam-se alcançar na pesquisa.

### 2.5.1.3 O algoritmo do BLAST

O BLAST é uma derivação do algoritmo Smith e Waterman (1981), que se caracteriza por apresentar a pontuação máxima do alinhamento local de duas sequências (PERTSEMLIDIS; FONDON, 2001). O algoritmo é capaz de identificar relação entre sequências que compartilham similaridade, mesmo que esta ocorra somente em algumas regiões isoladas. O algoritmo do BLAST divide a sequência em 11 (padrão word size) nucleotídeos e esses passam a ser procurados em todo o banco de dados. O algoritmo encontra trechos na sequência alvo e é utilizado como "*seeds*". A extensão do alinhamento é feita utilizando métodos para encontrar o alinhamento ótimo. Como o BLAST é uma ferramenta de alinhamento local, só será alinhado trechos que produzam um escore elevado.

De acordo com Altschul et al. (1990), o algoritmo do BLAST pode ser dividido em três estágios básicos: Compilação de uma lista de palavras de alta pontuação, procura destas palavras no banco de dados e extensão de alinhamentos a partir da palavras encontradas.

### 2.5.1.4 A estatística do BLAST - como avaliar um alinhamento

Alinhamentos locais também são passíveis de análise estatística rigorosa, que permite ao usuário decidir se um acerto é estatisticamente significativo. O algoritmo BLAST aproveita esses resultados rigorosos para minimizar o tempo gasto em alinhamentos que provavelmente não excedam a pontuação limite na qual ocorrem semelhanças aleatórias. Isso torna o BLAST muito mais rápido do que muitos outros algoritmos heurísticos. Um exemplo de alinhamento e seu respectivo sistema de pontuação pode ser observado nas Figuras 3 e 4.

Os traços verticais representam identidade (resíduos iguais) entre nucleotídeos, os traços horizontais representam *gaps*, o *query* representa a sequência submetida ao programa e o *subject* representa a sequência do banco de dados alinhada a sequência submetida.

As posições nas quais uma letra é pareada com um nulo é chamada de gap. Os gaps possuem scores negativos. Uma vez que eventos de mutação podem causar a inserção ou a deleção de mais de um resíduo, a presença de um *gap* é frequentemente considerada mais

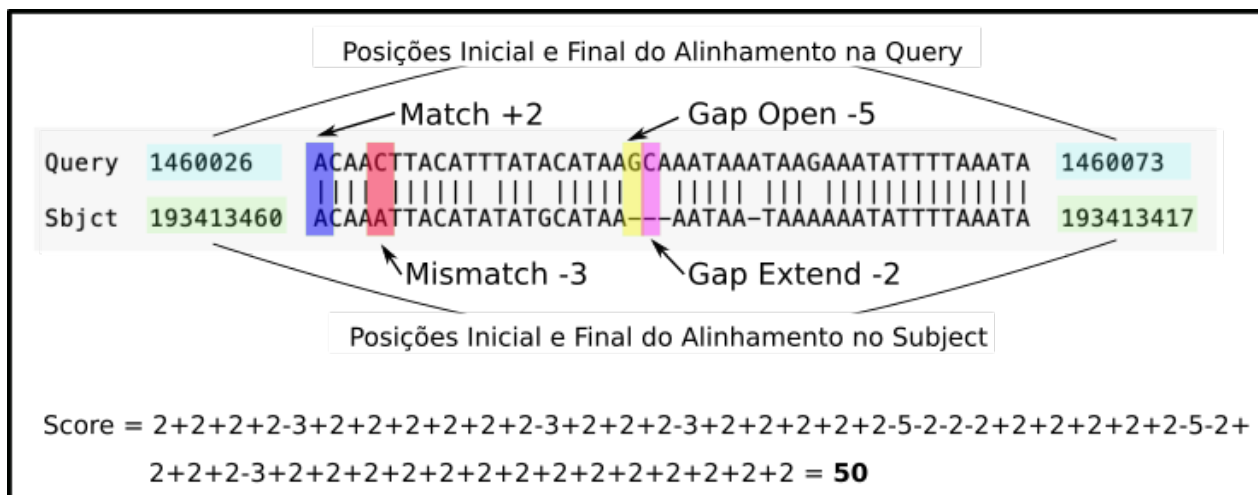


Figura 3: Representação da estatística de alinhamento feita pelo BLAST. O sistema de pontuação é regido pelas pontuações de *match*, *mismatch* e *gap*. O somatório destas pontuações é utilizado para calcular o *Score* de alinhamento.

significante do que o tamanho do *gap*. Portanto, um *gap* é penalizado com um grande peso, enquanto que cada resíduo subsequente no *gap* possui uma penalidade de peso menor.

A qualidade de cada par de sequências alinhadas é representado por um valor, o *score*. *Score* é o valor associado a um alinhamento, baseado em punições, relacionadas a *gaps* e a substituições, e em premiações, relacionadas a identidades. O melhor alinhamento entre duas sequências é aquele que maximiza o escore. O escore é calculado seguindo a Equação

$$Score = \#Matches * (+2) + \#Mismatch * (-3) + \#Gapopen * (-5) + \#Gap * (-2), \quad (2.1)$$

onde  $\#Matches$  significa quantidade de igualdade,  $\#Mismatch$  significa quantidade de desigualdade e  $\#Gap$  significa quantidade de lacunas. Quanto maior o score, melhor é o alinhamento, mas a significância do alinhamento não pode ser deduzida somente pelo score; deve-se considerar também outros fatores como o *e-value*.

O *match* indica uma correspondência entre o nucleotídeo da *query* com o nucleotídeo do *subject*, já o *mismatch* indica que não ocorreu correspondência entre nucleotídeos da *query* e do *subject*.

Os alinhamentos resultados do programa BLAST apresentam valores de pontuação e a esses valores é associado um valor chamado de *expected-value* (*e-value*). O *e-value* indica o número de alinhamentos que seriam esperados apresentando valores de *score* iguais ou melhores que o encontrado por acaso, dado o tamanho do banco de dados.

Não há um limiar perfeito para o *e-value*, pois depende muito do que se esteja trabalhando. O valor é basicamente uma medida de quantos alinhamentos você esperaria encontrar em um banco de dados deste tamanho por acaso. É dependente do comprimento

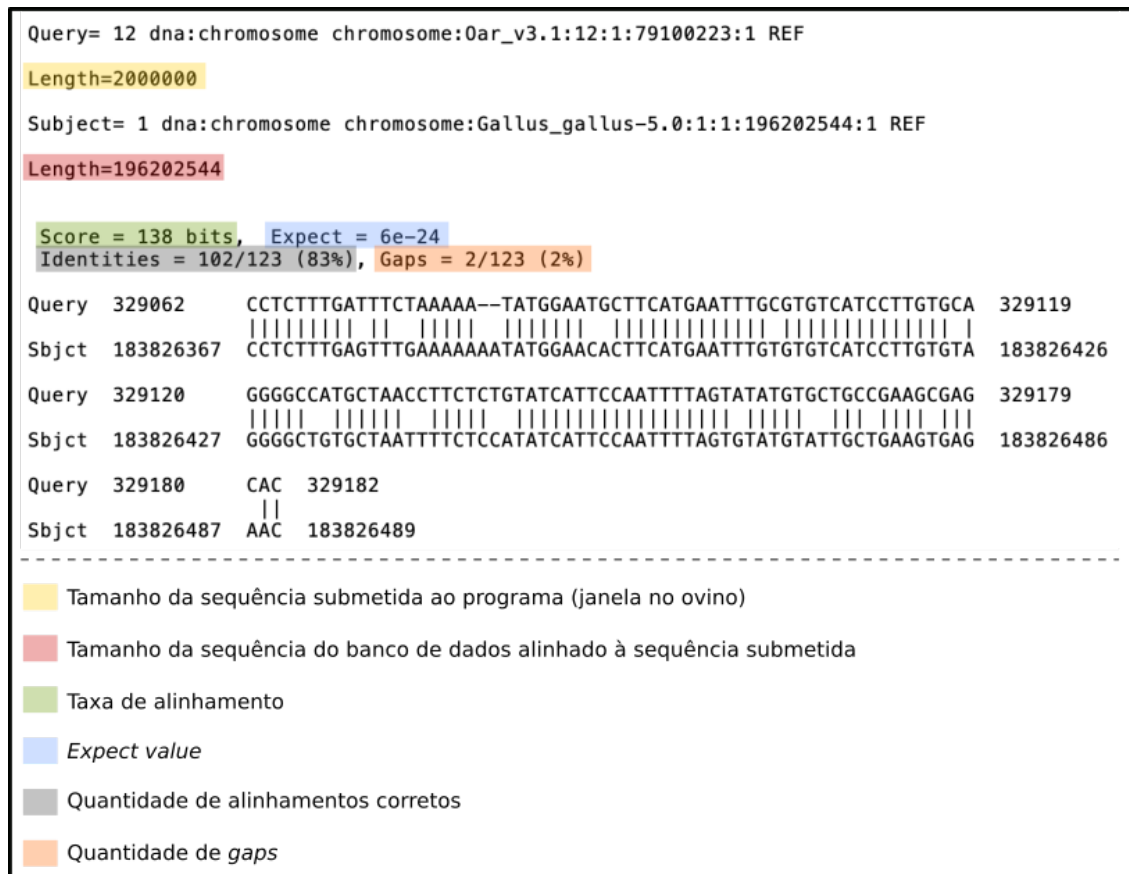


Figura 4: Saída do BLAST para o alinhamento entre duas sequências. Os *lengths* representam os tamanhos da *query* e *subject*. As pontuações de *match*, *mismatch* e *gaps*, irão determinar o *score* do alinhamento. A *Identities* é resultante da quantidade *matches* dentro do tamanho total do alinhamento. O *Expect* representa a confiabilidade do alinhamento.

do alinhamento, da pontuação de bit score do alinhamento e do tamanho do banco de dados que esta sendo usado na pesquisa.

$$E = \frac{m.n}{2^S}. \quad (2.2)$$

Quanto menor o valor E, ou quanto mais próximo de zero, mais significativa é a correspondência. No entanto, deve-se saber que alinhamentos curtos virtualmente idênticos têm valores E relativamente altos. Isso ocorre porque o cálculo do valor E leva em conta o comprimento da sequência de consulta. Esses altos valores de E fazem sentido porque sequências mais curtas têm maior probabilidade de ocorrer no banco de dados apenas por acaso.

#### 2.5.1.5 A sensibilidade do BLAST

Para comparações de sequências de nucleotídeos, o tamanho da palavra padrão (*word size*) é 11, onde nenhum par de segmentos pode ser marcado a menos que contenha



uma combinação perfeita de pelo menos 11 bases consecutivas. Pode-se ajustar o *word size*, por exemplo, diminuir o valor para aumentar a sensibilidade ao preço da velocidade.

## 2.6 Trabalhos relacionados com esta pesquisa

O estudo de genes candidatos associados a características de interesse econômico é uma ótima estratégia para a produção animal. São genes de ação biológica conhecida e que estão envolvidos com o desenvolvimento ou a fisiologia de uma característica de interesse econômico (P.F.BRYNE; M.D.MCMULLEN, 1996). Alguns exemplos bem sucedidos da aplicação desta estratégia são os genes halotano e RN, relacionados à qualidade de carne em suínos (VRIES *et al.*, 1998) e o da miostatina, associado à formação da musculatura dupla em bovinos Grobet *et al.* (1997). Marcadores para o gene ESR (receptor do estrogênio) em suínos foram associados ao número de leitões nascidos vivos e ao total de leitões nascidos vivos na primeira e última parição (ALFONSO, 2005). Um total de 13 SNPs foi associado a quatro genes relacionados à qualidade da carne e da carcaça em bovinos (HAEGEMAN *et al.*, 2003). Em ovinos, na raça Booroola Merino o gene BMPR-IB foi relacionado à fecundidade e nas raças Invedale e Hanna o gene BMP15 foi associado à ovulação (LIU; JIANG; DU, 2003). Outros exemplos podem ser encontrados para bovinos, suínos, aves e ovinos em (DEKKERS, 2004).

A principal limitação da aplicação da técnica de genes candidatos é que somente uma pequena proporção dos genes que controlam características quantitativas é conhecida. Dificuldades também existem no estabelecimento definitivo do efeito do gene candidato, pois a identificação da variante causal para um gene de efeito menor pode não ser facilmente determinada (COUTINHO *et al.*, 2007).



### 3 Genômica comparativa para identificação de genes candidatos à resistência a endoparasitas gastrintestinais na espécie *Ovis aries*

## Identificação de genes candidatos à resistência a endoparasitas gastrintestinais na espécie *Ovis aries*

Lilian R G Silva<sup>a</sup> e José L R Sarmiento<sup>a</sup>

<sup>a</sup>Programa de Pós-Graduação em Ciência Animal

---

**Resumo:** O controle de infecções helmínticas em ovinos tradicionalmente é feito com o uso de vermífugos, produto este que quando utilizado de forma indiscriminada pode levar a seleção de vermes resistentes. Neste sentido, é importante o desenvolvimento de formas alternativas de controle que minimizem o uso e gastos com anti-helmínticos na criação de ovinos. Como método de controle alternativo pode ser feita a identificação de ovinos que possuem genes de resistência a endoparasitas gastrintestinais para futura seleção genômica e melhoramento animal. Este trabalho teve por objetivo identificar no genoma da espécie *Ovis aries* genes e regiões candidatos à resistência a endoparasitas gastrintestinais por meio da análise e comparação genômica. Para identificação dos genes candidatos foram feitas buscas dentro de 6 janelas de 2Mb entorno de SNPs associados a fenótipos de resistência no genoma ovino, e posteriormente foram realizadas análises comparativas de DNA entre as janelas e genes das espécies *Capra hircus*, *Bos taurus*, *Sus scrofa*, *Mus musculus* e *Homo sapiens*. Para visualizar os genomas, verificar função de genes, comparar DNA e visualizar resultados, foram utilizados o Ensembl, UniProt, BLASTn e Kablammo, respectivamente. Os alinhamentos gerados pelo BLASTn passaram por filtros de identidade, RAG (Razão entre o tamanho do Alinhamento e tamanho do Gene) e *e-value*, onde foram identificados 281 alinhamentos significativos conforme esses filtros. A metodologia de investigação genômica norteadada pela informação de SNPs e da genômica comparativa possibilitaram a identificação dos genes FADD, LOC10111857, SLC22A18, ENSOARG0000000478, PRKN, CTSL, SLIT2, UBR3, KLHL41 e KLHL41, e de duas regiões localizadas respectivamente na posição 14658370-14659236 do cromossomo 12 e 46434090-46435478 do cromossomo 2. Ambos os genes e regiões passam a ser candidatos à resistência a endoparasitas gastrintestinais na espécie *Ovis aries*. As informações geradas nesta pesquisa servirão de base para que trabalhos mais aprofundados possam identificar os genes responsáveis pelo controle e expressão da resistência, que irão auxiliar na seleção e melhoramento de ovinos.

*Palavras-chave:* SNP. Alinhamentos genômicos. Parasitoses helmínticas. Ovinos.

---

## 3.1 Introdução

O Brasil tem uma longa tradição na criação de ovinos, com principal destaque para produção de carne, leite e lã. Segundo o [Instituto Brasileiro de Geografia e Estatística \(2018\)](#), no ano de 2017, o país apresentou um rebanho ovino de mais de 18 milhões de cabeças, com maior rebanho nas regiões Nordeste e Sul.

A produção de ovinos encontra-se em expansão devido ao aumento do consumo de carne destes animais. No entanto, entraves na produção, como infecções parasitárias gastrintestinais, tem comprometido o aproveitamento econômico destes animais. Para tanto os programas de seleção e melhoramento genético de ovinos devem incluir além de características relacionadas à produção, a resistência a endoparasitas ([ALBERS et al., 1987](#); [AMARANTE et al., 2004](#)).

As infecções causadas por nematódeos gastrintestinais são uma das doenças parasitárias mais prevalentes que afetam pequenos ruminantes no mundo todo ([TORRES-ACOSTA; HOSTE, 2008](#)). Segundo [Melo et al. \(2003\)](#) e [Nova et al. \(2014\)](#) as helmintoses provocam grandes prejuízos, como, queda no desempenho dos animais, elevação dos custos de produção, e até mesmo a morte dos animais. Segundo [Waller et al. \(1996\)](#), o controle tradicional da doença é feito com o uso de drogas anti-helmínticas, porém o uso indiscriminado das mesmas podem levar a seleção de vermes resistentes, contaminação meio ambiente e contaminação dos produtos oriundos dessa cadeia produtiva. Além do que, a compra de vermífugos gera altos gastos para o produtor.

Uma forma alternativa para diminuir o uso de anti-helmínticos consiste em conhecer o material genético do animal de interesse, ou seja, identificar se o ovino possui genes que promovam maior ou menor resistência a endoparasitas gastrintestinais. De acordo com [Daetwyler et al. \(2010\)](#), a identificação e estudo de marcadores moleculares tipo SNP (*Single Nucleotide polymorphism*) juntamente com estudos de Associação Genômica Ampla (GWAS - *Genome Wide Associations Study*), permitem identificar regiões no genoma associados com características de interesse. Os estudos de GWAS auxiliam no entendimento dos mecanismos genéticos que controlam características de herança complexa, como a resistência genética a verminose. A partir disso pode-se analisar se próximo a essas regiões de interesse há genes candidatos a característica.

De acordo com [BIAGIOTTI \(2016\)](#), os estudos de associação genômica tem por objetivo identificar variantes genéticas em um grande número de indivíduos, a fim de se descobrir associações estatísticas significativas com fenótipos de interesse. A associação de marcadores moleculares com características de interesse econômico, abrem caminhos para que se possa analisar regiões próximas ao marcador e identificar genes candidatos à expressão e controle da característica. Porém, pode ocorrer que próximo ao SNP associado existam poucas ou nenhuma informação sobre genes no genoma da espécie. Para contribuir

com a busca de genes para essas regiões, pode-se lançar mão da genômica comparativa, que permite a comparação do material genético de diferentes espécies, e a extrapolação de informações entre as mesmas.

Os avanços ocorridos nas áreas da biologia molecular, genômica, estatística e ciências da computação, permitiram a criação da bioinformática, campo que alavancou e otimizou os estudos na área da genômica. Através da bioinformática criou-se ferramentas computacionais com grande poder de análise de dados volumosos advindos do sequenciamento genômico. A exemplo, pode-se destacar a ferramenta de busca e alinhamento local BLAST (*Basic Local Alignment Search Tool*). A ferramenta BLAST possibilita a análise de um grande volume de dados que podem ser usados, por exemplo, na criação de mapas físicos e genéticos. Dessa forma, a comparação entre mapas possibilita a extrapolação de informações de uma espécie para outra através de alinhamentos genômicos. O alinhamento permite estabelecer uma correspondência entre duas sequências de DNA (HAGEN, 2001; DEROME et al., 2018).

A partir da problemática apresentada, objetivou-se com este trabalho, identificar genes candidatos à resistência a endoparasitas gastrintestinais em ovinos, analisando regiões específicas de DNA da espécie *Ovis aries*, e compara-las com genes descritos nos genomas das espécies *Capra hircus*, *Bos taurus*, *Sus scrofa*, *Mus musculus* e *Homo sapiens*.

## 3.2 Metodologia

Esta pesquisa teve como base o trabalho de associação e seleção genômica ampla em ovinos Santa Inês, realizado por Biagiotti (BIAGIOTTI, 2016). O trabalho teve por objetivo identificar SNPs associados com fenótipos de resistência a endoparasitoses em ovino. Foram utilizados fenótipos de presença de pelo arrepiado, presença de ovos do gênero *Strongylus* nas fezes, OPG (Ovos por grama de fezes), coloração da mucosa (FAMACHA), parâmetros sanguíneos, contagem diferencial de células e escore de condição corporal. Foram analisados 271 ovinos com registro na Associação Brasileira de Criadores de Ovinos. Os animais foram genotipados com SNPChip *Ovine* de 50K. A associação genômica ampla foi realizada por regressão corrigida para efeitos aleatórios poligênicos visando associação entre o marcador e o fenótipo. Como resultado do trabalho foram identificados 6 SNPs com associação significativa, como mostrado na Tabela 3.

A Tabela 3 contém o nome de cada SNP, posição no cromossomo ovino e fenótipo associado ao marcador. Todas estas informações foram o ponto de partida desta pesquisa, para nortear a busca e identificação de genes candidatos à resistência em ovinos.

Tabela 3: Lista de SNPs com suas respectivas posições no cromossomo do *Ovis aries* e fenótipos de resistência ao qual foram associados.

SNP	Posição do SNP	Cromossomo	Característica
OAR2_148350187.1	139613559	2	LogOPG
OAR8_91253490_X.1	84523257	8	Presença de ovos do gênero <i>Strongylus</i>
OAR12_17005059.1	14286956	12	Presença de ovos do gênero <i>Strongylus</i>
OAR21_52583090.1	47316667	21	Pelo arrepiado
s23985.1	31883314	2	Escore condição corporal
s26633.1	39230785	6	LogOPG

Adaptado de [BIAGIOTTI \(2016\)](#) com as posições atualizadas pela Illumina (OvineSNP50).

### 3.2.1 Análise de janelas para a identificação de genes candidatos a resistência

A busca por genes candidatos iniciou-se com a escolha e visualização do genoma da espécie *O. aries* no navegador Ensembl. Após seleção do genoma, foi inserido a localização dos SNPs descritos na Tabela 3. Entorno de cada um dos 6 SNPs, foi analisada uma região de 1Mb à esquerda e 1Mb à direita do marcador (janela de 2Mb). A representação esquemática desta etapa consta na Figura 5.

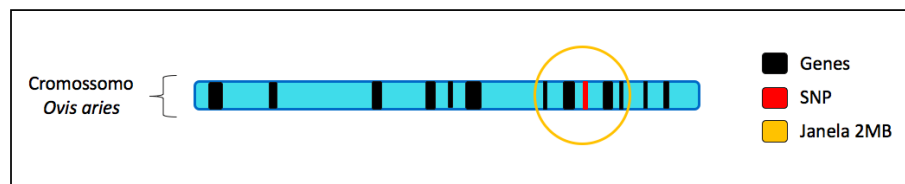


Figura 5: Localização do SNP no cromossomo e delimitação da janela para análise de genômica comparativa.

Dentro de cada uma das 6 janelas de 2Mb buscou-se por genes codificantes de proteína. A função biológica de cada gene foi analisada no banco de dados UniProt, sendo que, os genes que apresentaram relação com processos imunológicos foram selecionados e anotados. Por fim, a primeira etapa da pesquisa foi concluída com a escolha de todos os genes envolvidos com processos imunológicos, como candidatos à resistência a parasitas gastrintestinais em ovinos.

### 3.2.2 Genômica comparativa para identificação de genes candidatos a resistência

A genômica comparativa, fez-se necessária para identificar genes relacionados com processos imunológicos em outras espécies, que possuíssem similaridade com segmentos

das janelas de 2Mb, e pudessem extrapolar informações para o genoma da espécie *O. aries*. A Figura 6 representa a metodologia completa da segunda etapa da pesquisa.

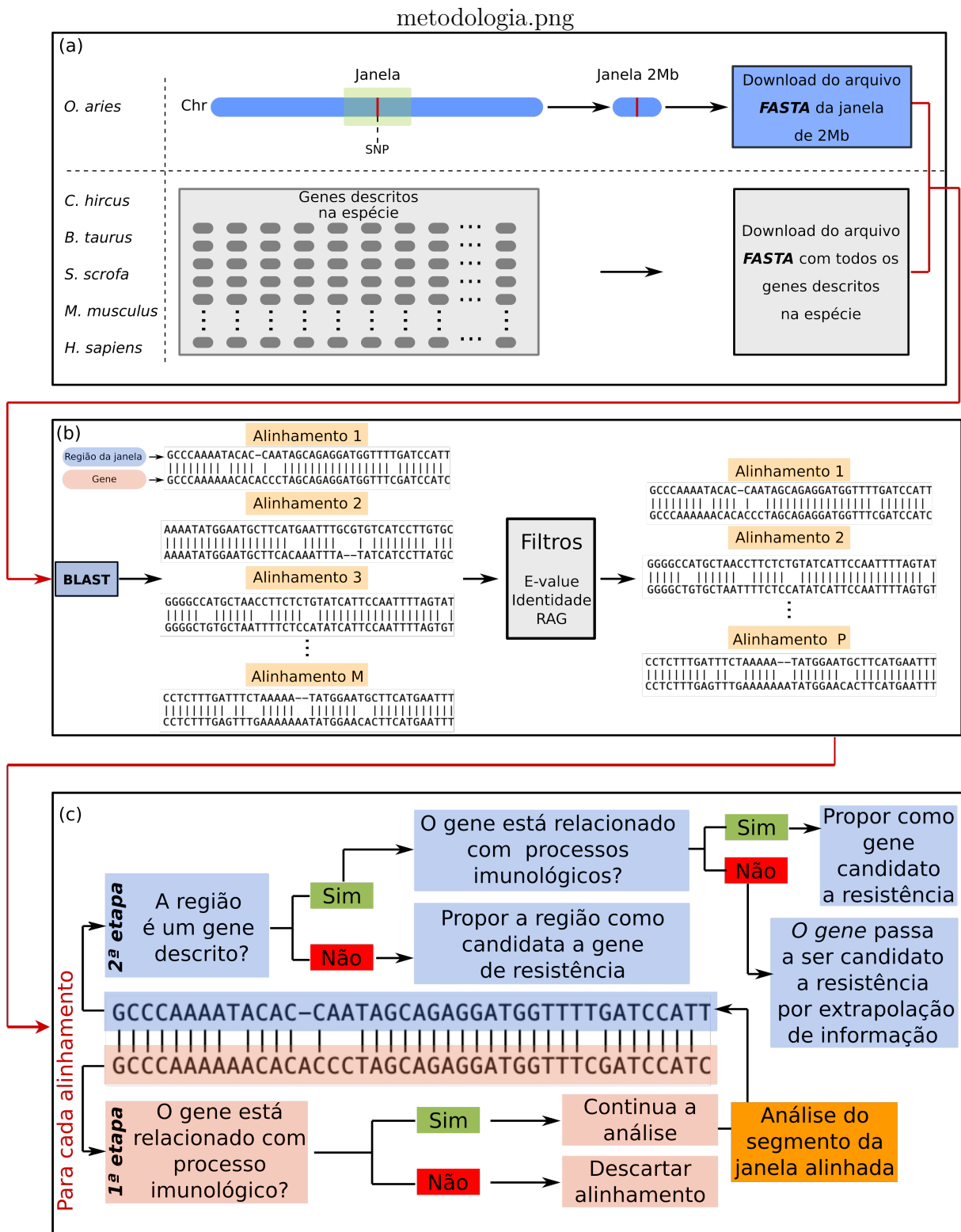


Figura 6: Metodologia completa do trabalho: a) Etapa de seleção dos seguimentos a serem alinhados. b) Alinhamentos gerados pelo BLASTn. c) Análise dos genes oriundos dos melhores alinhamentos.

A segunda etapa da pesquisa consistiu em comparar as 6 janelas do *O. aries* com os genes descritos nas espécies *Capra hircus*, *Bos taurus*, *Sus scrofa*, *Mus musculus* e *Homo sapiens*. Para comparar as sequências de nucleotídeos, foi utilizado o programa BLASTn com a tarefa BLASTn, específico para este tipo de análise, e que exige uma correspondência exata de 11 nucleotídeos para iniciar o alinhamento.

Todo o material a ser comparado precisou está no formato de arquivo FASTA, pois este é compatível com a entrada do BLAST. Baixou-se no Ensembl o arquivo FASTA correspondente de cada janela e de cada conjunto gênico das espécies a serem comparadas. Os arquivos foram submetidos ao BLASTn, onde o algoritmo realizou o alinhamento de duas sequências de DNA por vez. O algoritmo buscou nas janelas do *Ovis aries* regiões que tivessem similaridade com os genes das espécies *Capra hircus*, *Bos taurus*, *Sus scrofa*, *Mus musculus* e *Homo sapiens*.

O BLASTn gerou muitos alinhamentos, devido a alta sensibilidade do algoritmo em detectar 11 nucleotídeos para iniciar um alinhamento, como mencionado anteriormente. A sensibilidade alta faz com que o alinhamento seja lento, porém mais confiável. Por conta da quantidade de alinhamentos gerados, os mesmos precisaram passar por filtros a fim de encontrar os alinhamentos mais significativos.

Após a filtragem restaram apenas os alinhamentos significativos, onde se pode fazer a análise das sequências de DNA, estando esta dividida em duas etapas. A primeira etapa consistiu em verificar no bancos de dados UniProt, se o gene que alinhou com a janela possuía relação com processos imunológicos, caso contrário o alinhamento era descartado.

A segunda etapa consistiu em avaliar a região da janela do *O. aries* que alinhou com o gene da outra espécie. Se a região da janela for um gene relacionado com processos imunológicos, o mesmo passa a ser candidato à resistência. Caso a região da janela for um gene não relacionado com processos imunológicos, extrapola-se a informação do gene da outra espécie para o gene ovino. Entretanto, se a região da janela ainda não for descrita como gene, a mesma passa a ser candidata a gene de resistência, mediante a extrapolação de informação do gene da outra espécie.

### 3.2.3 Filtragem dos alinhamentos gerados pelo BLASTn *E-Value*

Com base no objetivo do problema (encontrar similaridades entre espécies), utilizou-se a métrica de alinhamento *E-Value*(*EV*) como primeiro passo na filtragem dos alinhamentos. Utilizamos  $EV = 1e - 100$ , ou seja, todos os alinhamentos com *EV* maior que  $1e - 100$  foram descartados. Tal *EV* irá retornar alinhamentos que não foram obtidos ao acaso (WAN et al., 2014).

### 3.2.4 Filtragem por Identidade

Visto que para cada comparação feita pelo BLASTn (genes das 5 espécies alinhados com as 6 janelas do ovino) foram obtidos uma elevada quantidade de alinhamentos mesmo após a aplicação do filtro por E-value, utilizou-se apenas o melhor alinhamento de cada gene.

Para realizar essa filtragem, utilizou-se o valor da identidade do alinhamento como métrica, pois ela reflete a qualidade dos alinhamentos obtidos. Dessa forma, dentre todos os alinhamento obtidos para cada gene, foi escolhido aquele com maior valor de identidade para ser levado adiante.

### 3.2.5 Filtragem pela Razão entre o Tamanho do Alinhamento e o Tamanho do Gene (RAG)

Utilizou-se a Equação 3.1 para realizar a última etapa de filtragem dos alinhamentos. Para avaliar se houve o máximo de alinhamento entre o tamanho total do gene e o segmento da janela, utilizou-se a razão entre os tamanhos do alinhamento e do gene. O objetivo foi identificar a proporção do gene que alinhou. Por exemplo, caso um gene possua 1000 pb e em determinado alinhamento, somente 600 pb estejam presentes, então  $RAG = 600/1000 = 0,6$ . Nesse caso, 60% do gene alinhou com alguma região do *O. aries*. Dessa forma, quanto mais próximo de 1 for essa razão mais similar são as duas sequências. Foi utilizada 0,9 como limiar para o RAG, ou seja, somente os alinhamentos que tiverem mais que 90% de similaridade com o gene que alinhou foram levados adiante.

$$RAG = \frac{\text{Tamanho do Alinhamento}}{\text{Tamanho do Gene}}. \quad (3.1)$$

## 3.3 Resultado e discussão

Os resultados foram divididos em 4 etapas: 1) Genes candidatos à resistência identificados no genoma do *O. aries*; 2) Alinhamentos significativos do BLASTn; 3) Identificação de genes na janela do *O. aries* que alinharam com genes das espécies comparadas; 4) Identificação de genes candidatos à resistência via genômica comparativa.

### 3.3.1 Genes candidatos à resistência identificados no genoma do *O. aries*

A partir da investigação feita próximo aos 6 SNPs, foram identificados um total de 56 genes codificantes de proteínas, dos quais apenas 10 apresentaram função biológica relacionada com processos imunológicos. Os 10 genes foram selecionados como candidatos à resistência a endoparasitas gastrintestinais na espécie *O. aries*.



A Tabela 4 traz as informações sobre o gene, em qual janela de SNP o gene foi encontrado no genoma *O. aries*, e a localização do SNP dentro da janela. Dados importantes para que se possa encontrar de forma mais rápida os genes de interesse.

Tabela 4: Genes candidatos à resistência na espécie *O. aries*, e suas respectivas localizações em cada janela.

Gene	SNP da janela	Localização dos candidatos
FADD	OAR21_52583090.1	Chr21 46672617-46674854
LOC101118574	OAR21_52583090.1	Chr21 47631203-47632150
SLC22A18	OAR21_52583090.1	Chr21 47911171-47935403
ENSOARG00000004781	OAR8_91253490_X.1	Chr8 83627065-83633679
PRKN	OAR8_91253490_X.1	Chr8 83818709-84599387
CTSL	s23985.1	Chr2 31785406-31790389
SLIT2	s26633.1	Chr6 39939556-40091549
UBR3	OAR2_148350187.1	Chr2 138424294-138625167
KLHL41	OAR2_148350187.1	Chr2 138842626-138856360
STK39	OAR2_148350187.1	Chr2 140068741-140315412

A função biológica e notação dos genes candidatos, de acordo com o banco de dados UniProt ([CONSORTIUM, 2018](#)), estão listados a seguir:

- FADD esta relacionado com molécula adaptadora apoptótica que recruta caspase-8 ou caspase-10 para os receptores ativados Fas (CD95) ou TNFR-1. Possui 2 éxons de codificação, e comprimento de transcrição de 633 pb.
- LOC101118574 permite a atividade do receptor acoplado à proteína G, e está envolvido na via de sinalização do receptor acoplado à proteína G. Possui 1 éxon de codificação, e comprimento de transcrição de 948 pb.
- SLC22A18 permite a atividade transportadora transmembrana de drogas, é um componente integral da membrana, e permite a ligação da proteína ligase da ubiquitina. Possui 12 éxons de codificação, e comprimento de transcrição de 4208 pb.
- ENSOARG00000004781 permite a atividade ubiquitina-proteína transferase. Possui 2 éxons de codificação, e comprimento de transcrição de 240 pb.
- PRKN permite a ligação da proteína ligase da ubiquitina, e envolvido na proteína de ubiquitinação. Possui 10 éxons de codificação, e comprimento de transcrição de 1113 pb.
- CTSL está envolvido na resposta imune adaptativa. Possui 7 éxons de codificação, e comprimento de transcrição de 1,026 pb.
- SLIT2 está envolvido com a regulação negativa da via de sinalização mediada por quimiocinas. Possui 33 éxons de codificação, e comprimento de transcrição de 4210 pb.

- UBR3 está envolvido com proteína de ubiquitinação, permite a atividade da proteína ligase ubiquitina, permite atividade ubiquitina-proteína transferase, e permite atividade da proteína ligase ubiquitina. Possui 41 éxons de codificação, e comprimento de transcrição de 7062 pb.
- KLHL41 faz parte do complexo de ubiquitina ligase Cul3-RING, e está envolvido com a proteína de ubiquitinação. Possui 6 éxons de codificação, e comprimento de transcrição de 1821 pb.
- STK 39 está envolvido com a resposta celular à quimiocina. As quimiocinas e seus receptores são capazes de controlar a migração e a residência de células imunes. Possui 13 éxons de codificação, e comprimento de transcrição de 2417 pb.

### 3.3.2 Alinhamentos ótimos do BLASTn

Após a filtragem de todos os alinhamentos gerados entre as janelas do ovino e as espécies *Capra hircus*, *Bos taurus*, *Sus scrofa*, *Mus musculus* e *Homo sapiens*, chegou-se ao número de 281 alinhamentos ótimos (Tabela 5). Todos os genes dos 281 alinhamentos tiveram sua função biológica analisada no banco de dados UniProt.

Tabela 5: Quantidade de alinhamentos por janela antes e após os filtros de *E-value*, Identidade e RAG (Razão entre o tamanho do Alinhamento e tamanho do Gene).

Janela	Alinhamentos	Alinhamentos após Filtros		
		<i>E-Value</i>	Identidade	RAG
OAR2_148350187.1	58310	2228	222	44
OAR8_91253490_X.1	65492	1590	105	28
OAR12_17005059.1	60687	3663	166	27
OAR21_52583090.1	41903	1387	222	57
s23985.1	58488	2776	208	62
s26633.1	61219	2645	176	63
<b>Total</b>	<b>346099</b>	-	-	<b>281</b>

O BLASTn gerou um total de 346099 alinhamentos, que passaram por filtros de *E-value*, Identidade e RAG (Razão entre o Tamanho do Alinhamento e o Tamanho do Gene), restando um total de 281 alinhamentos significativos. Estes alinhamentos representam os genes das espécies comparadas que alinharam com segmentos das janelas. e que posteriormente foram analisados.

### 3.3.3 Identificação de genes nas janelas do *O. aries* que alinharam com genes das espécies comparadas

A partir da análise dos alinhamentos ótimos gerados pelo BLASTn, pode-se identificar que genes presentes nas janelas do *O. aries* alinharam com genes das espécies *Capra*

*hircus*, *Bos taurus*, *Sus scrofa*, *Mus musculus* e *Homo sapiens*, como se observa nas Figuras 7 e 8.

Os alinhamentos entre genes (7 e 8), gerados com ferramenta a Kablammo (em escalas de pares de base (pb)), de acordo com o filtro RAG=0,9 (alinhamento de 90% a 100% entre genes), tiveram suas funções analisadas no banco de dados UniProt, para que pudessem ser selecionados como genes importantes apenas os que possuíssem relação com processos imunológicos.

Dos 6 alinhamentos presentes na Figura 7, apenas o alinhamento entre o gene LOC102159152 do *Sus scrofa* e o gene LOC101118574 do *Ovis aries* (Figura 7a)) foi relevante quanto ao fato de haver função biológica descrita para ambos e estarem relacionados com processos imunológicos. A análise da função biológica revelou que os dois genes permitem a atividade do receptor acoplado à proteína G, e estão envolvidos na via de sinalização do receptor acoplado à proteína G. A proteína G pertence a uma classe de proteínas envolvidas na transdução de sinais celulares (CONSORTIUM, 2018). Esse compartilhamento de informações reforça que o gene LOC101118574 do *Ovis aries* é candidato à resistência a parasitas gastrintestinais.

Os demais alinhamentos presentes na Figura 7 foram descartados, visto que a anotação biológica não possui função descrita ou não está relacionada com processos imunológicos, o que sugere que suas funções estão associadas a outras características ainda não estudadas.

Na Figura 8 também há a representação gráfica de mais 6 alinhamentos ocorridos entre genes do *O. aries* e os genes das espécies comparadas. Ao analisar a função biológica no UniProt, verificou-se que nenhum gene apresenta relação com processos imunológicos. Sendo assim, não pode-se candidatar nenhum gene à resistência a endoparasitas gastrintestinais na espécie *O. aries*.

### 3.3.3.1 Identificação de genes candidatos à resistência via genômica comparativa

Dentro dos 281 alinhamentos ótimos, houve similaridade entre 4 genes e dois segmentos de janelas 3.3.3.1. Os genes pertencem as espécies *M. musculus*, *B. taurus* e *H. sapiens*, e estão relacionados com processos imunológicos, fato verificado no banco de dados UniProt (CONSORTIUM, 2018). Os dois segmentos de DNA das janelas do *O. aries* não estão descritos como genes, ou seja, são regiões que precisam de estudos mais aprofundados. Por conta da alta similaridade demonstrada mediante alinhamento genômico estas duas regiões passam a ser candidatas a genes, e mediante a extrapolação de informações do genes das outras espécies que alinharam, essas regiões passam a ser candidatas a genes de resistência.

Como pode-se observar na Figura 3.3.3.1, os genes EEF1A1 (*B. taurus*), Eef1a1 (*M.*

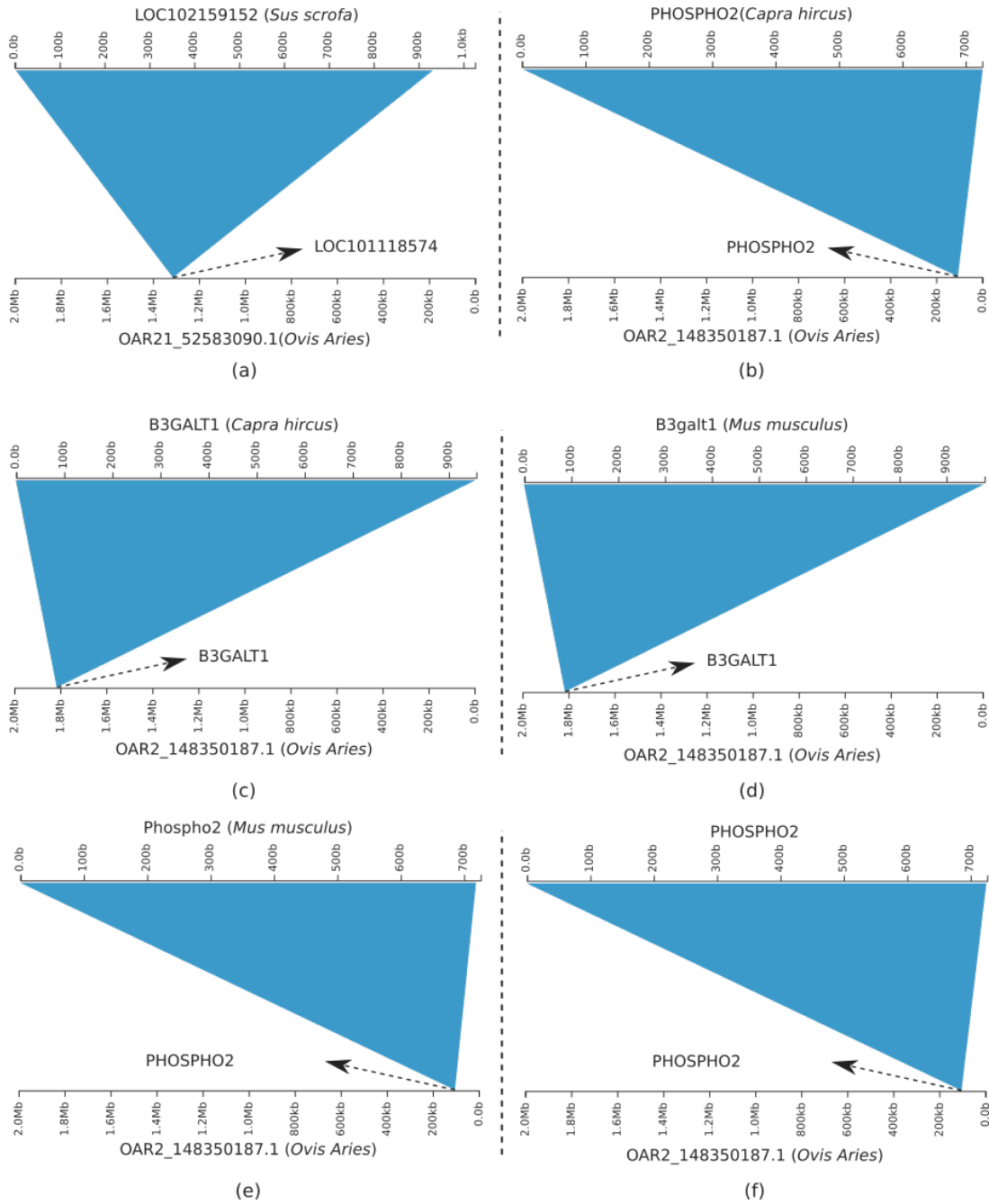


Figura 7: Representação gráfica em escalas de pares de base (pb), de acordo com o filtro RAG=0,9, do alinhamento ocorrido entre genes das janelas e genes das espécies comparadas.

Tabela 6: Posições dos genes e dos segmentos das janelas que obtiveram similaridade identificada.

Espécies comparadas		Região candidata a gene de resistência no <i>O. aries</i>
Gene (Espécie)	Posição no cromossomo	Posição no cromossomo
Rpsa ( <i>M. musculus</i> )	9:120127689-120132369	12:14658370-14659236
EEF1A1 ( <i>B. taurus</i> )	9:13011231-13020524	2:46434090-46435478
Eef1a1 ( <i>M. musculus</i> )	9:78478449-78489151	2:46434090-46435478
EEF1A1 ( <i>H. sapiens</i> )	6:73515750-73523797	2:46434090-46435478

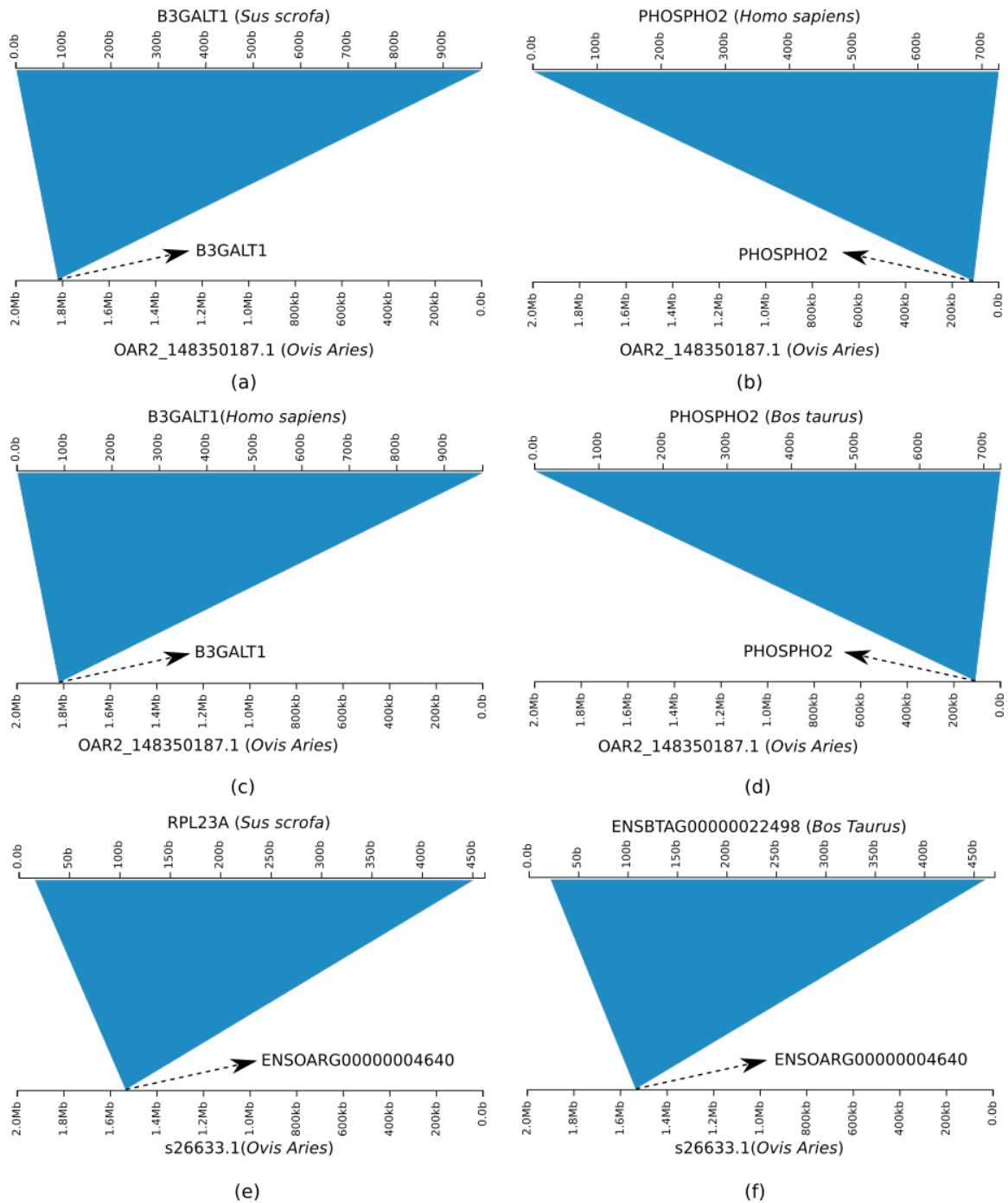


Figura 8: Representação gráfica em escalas de pares de base (pb), de acordo com o filtro RAG=0,9, do alinhamento ocorrido entre genes das janelas e genes das espécies comparadas.

*musculus*) e **EEF1A1** (*H. sapiens*) possuem similaridade com a mesma região do cromossomo 2 da espécie *O. aries*. Diferente dos demais genes, o gene **Rpsa** (*M. musculus*), obteve similaridade com uma região do cromossomo 12 da espécie *O. aries*. Os genes tiveram suas informações extrapoladas para as regiões do *O. aries*, que agora passam a ser regiões candidatas a resistência a parasitas gastrintestinais. A função dos genes alinhados com segmentos de DNA do *O. aries* estão descritas a seguir conforme informações do banco de dados UniProt ([CONSORTIUM, 2018](#)):

- Gene **Rpsa**, pertence a espécie *Mus musculus*, sua função é ativar células T citotóxicas precursoras do timo anti-OFA / iLRP específicas. Pode induzir células supressoras CD8 T secretoras de IL-10. O linfócito T-citotóxico apresenta receptores TCR. Especializado para o reconhecimento de antígenos associados ao complexo MHC-I na superfície de outras células. Produz perforinas e outras proteínas que matam células estranhas, células infectadas por vírus e algumas células cancerosas. O gene possui comprimento de 1938 pb, tem 6 transcritos (variantes de *splicing*) e 166 ortólogos.
- Gene **EEF1A1**, pertencente a espécie *Bos taurus*, sua função envolve as proteínas PARP-1 e TXK, onde juntas formam um complexo que atua como um fator de transcrição e especificação de células T helper 1 (Th1) e liga-se ao promotor de IFN-gama para regular diretamente a sua transcrição e estão, assim envolvidos de forma importante na produção de citocinas Th1. Este gene tem 2 transcritos (variantes de *splicing*), 4839 pb de comprimento, 193 ortólogos e 27 parálogos.
- gene **Eef1a1** da espécie *Mus musculos*, sua função envolve as proteínas PARP-1 e TXK, onde juntas formam um complexo que atua como um fator de transcrição e especificação de células T helper 1 (Th1) e liga-se ao promotor de IFN-gama para regular diretamente a sua transcrição e estão, assim envolvidos de forma importante na produção de citocinas Th1. Este gene tem 794 pb de comprimento, 8 transcritos (variantes de *splicing*), 301 ortólogos e 18 parálogos.
- gene **EEF1A1** da espécie *Homo sapiens*, a proteína deste gene promove a ligação de aminoacil-ARNt dependente de GTP ao sítio A dos ribossomas durante a biossíntese de proteínas. Com PARP1 e TXK, forma um complexo que atua como um fator de transcrição específico da célula T helper 1 (Th1) e se liga ao promotor do IFN-gama para regular diretamente sua transcrição, e está, portanto, envolvido de maneira importante na produção de citocinas Th1. Este gene tem 4441 pb de comprimento, 11 transcritos (variantes de *splicing*), 195 ortólogos e 18 parálogos.

Como resultado final da pesquisa constatou-se que nas janelas do genoma *O. aries* foram identificados 10 genes candidatos à resistência a endoparasitas gastrintestinais, e

via genômica comparativa foram identificadas duas regiões do *O. aries* candidatas a genes de resistência a parasitas gastrintestinais.

Pode-se constatar também, que a metodologia usada com base na investigação de regiões próximas a SNPs associados a genótipos de resistência e uso do BLASTn para comparar regiões de interesse foi válida, chegando-se a resultados consoantes com o objetivo da pesquisa.

### 3.4 Conclusão

A metodologia de investigação genômica norteada pela informação de SNPs e da genômica comparativa possibilitaram a identificação dos genes FADD, LOC10111857, SLC22A18, ENSOARG0000000478, PRKN, CTSL, SLIT2, UBR3, KLHL41 e KLHL41, e de duas regiões localizadas respectivamente na posição 14658370-14659236 do cromossomo 12 e 46434090-46435478 do cromossomo 2. Ambos os genes e regiões passam a ser candidatos a resistência a endoparasitas gastrintestinais na espécie *Ovis aries*.

As informações geradas nesta pesquisa servirão de base para que trabalhos mais aprofundados possam identificar os genes responsáveis pelo controle e expressão da resistência, que irão auxiliar na seleção e melhoramento de ovinos.

# Referências

- AL-SAMARAI, F.; AL-KAZAZ, A. A. Molecular markers: an introduction and applications. *European Journal of Molecular Biotechnology*, v. 9, p. 118–130, 2015. ISSN 2409-1332. Citado 2 vezes nas páginas 16 e 17.
- ALBERS, G. A. A. et al. The genetics of resistance and resilience to haemonchus contortus infection in young merino sheep. *International Journal for Parasitology*, v. 17, n. 7, p. 1355–1363, 1987. ISSN 1879-0135. Citado 2 vezes nas páginas 12 e 33.
- ALFONSO, L. *Use of meta-analysis to combine candidate gene association to study the relationship between the ESR PvuII polymorphism and sow litter size*. [S.l.: s.n.], 2005. 417-35 p. (Genetics, selection, evolution : GSE, v. 37). ISSN 0999-193X. Citado na página 30.
- ALTSCHUL, S. F. et al. Basic local alignment search tool. *Journal of Molecular Biology*, v. 215, n. 3, p. 403 – 410, 1990. ISSN 0022-2836. Citado na página 27.
- AMARANTE, A. F. T. et al. Resistance of santa ines, suffolk and ile de france sheep to naturally acquired gastrointestinal nematode infections. *Veterinary Parasitology*, v. 120, n. 1-2, p. 91–106, 2004. ISSN 1873-2550. Citado 3 vezes nas páginas 12, 16 e 33.
- AMARANTE, A. F. T. do. *Os parasitas de ovinos [online]*. [S.l.]: UNESP, 2014. ISBN 978-85-68334-42-3. Citado 2 vezes nas páginas 15 e 16.
- ANTHONY, R. M. et al. Protective immune mechanisms in helminth infection. *Nature Reviews Immunology*, v. 7, p. 975–987, 2007. ISSN 1474-1741. Citado na página 15.
- BADER, J. S. The relative power of snps and haplotype as genetic markers for association tests. *Pharmacogenomics*, v. 2, n. 1, p. 11–24, 2001. ISSN 1744-8042. Citado na página 18.
- BALIC, A.; BOWLES, V.; MEEUSEN, E. The immunobiology of gastrointestinal nematode infections in ruminants. *Advances in parasitology*, v. 45, p. 181–241, 2000. ISSN 0065-308X. Citado na página 15.
- BIAGIOTTI, D. *Associação e seleção genômica ampla em ovinos santa inês para características relacionadas a resistência à endoparasitas*. Tese (Doutorado) — Departamento de Zootecnia, Universidade Federal do Piauí, 2016. Citado 4 vezes nas páginas 12, 33, 34 e 35.
- BISHOP, S.; MORRIS, C. Genetics of disease resistance in sheep and goats. *Small Ruminant Research*, v. 70, n. 1, p. 48 – 59, 2007. ISSN 0921-4488. Special Issue: The Outlook of Quantitative and Molecular Genetics Applications in Improving Sheep and Goats. Citado na página 16.
- BISHOP, S.; STEAR, M. Modeling of host genetics and resistance to infectious diseases: understanding and controlling nematode infections. *Veterinary Parasitology*, v. 115, n. 2, p. 147 – 166, 2003. ISSN 0304-4017. Citado na página 15.



- BISHOP, S. C.; WOOLLIAMS, J. A. Genetic approaches and technologies for improving the sustainability of livestock production. *Journal of the Science of Food and Agriculture*, v. 84, n. 9, p. 911–919, 2004. ISSN 1097-0010. Citado 2 vezes nas páginas 15 e 16.
- BOWMAN, J.; SANDERS, M. *Análise Genética: Uma abordagem integrada*. 1a edição. ed. [S.l.]: Pearson Brasil, 2014. ISBN 9788543005911. Citado na página 22.
- BROOKES, A. J. The essence of snps. *Gene*, v. 234, n. 2, p. 177 – 186, 1999. ISSN 0378-1119. Citado na página 17.
- CARVALHO, P. C. F. et al. Importância da estrutura da pastagem na ingestão e seleção de dietas pelo animal em pastejo. In: *Anais da XXXVIII Reunião Anual da Sociedade Brasileira de Zootecnia*. [S.l.: s.n.], 2001. p. 853–871. Citado na página 15.
- CLARK, M. T. The next war: By caspar weinberger and peter schweizer. (washington d.c.: Regnery, 1996. 470 pp. \$27.50. *Orbis*, v. 42, n. 1, p. 121 – 130, 1998. ISSN 0030-4387. Citado na página 23.
- CONSORTIUM, T. U. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Research*, v. 47, n. D1, p. D506–D515, 2018. ISSN 0305-1048. Citado 3 vezes nas páginas 39, 41 e 44.
- COUTINHO, L. et al. *Genômica Animal*. 2007. <[http://www.cnpsa.embrapa.br/genomafrango/publica/20062007Zootec\\_Luiz%Coutinho\\_final2007.pdf](http://www.cnpsa.embrapa.br/genomafrango/publica/20062007Zootec_Luiz%Coutinho_final2007.pdf)>. Online; acessado em 03 Janeiro 2019. Citado 2 vezes nas páginas 18 e 30.
- COUTINHO, L. L.; ROSÁRIO, M. F. d.; JORGE, E. C. Biotecnologia animal. *Estudos Avançados*, Scielo, v. 24, p. 123 – 147, 2010. ISSN 0103-4014. Citado na página 18.
- D., W. *Immunity to parasites: how parasitic infections are controlled*. 2a edição. ed. [S.l.]: Cambridge University Press, 1992. ISBN 0521562457. Citado na página 14.
- DAETWYLER, H. D. et al. The impact of genetic architecture on genome-wide evaluation methods. *Genetics*, Genetics, v. 185, n. 3, p. 1021–1031, 2010. ISSN 0016-6731. Citado 2 vezes nas páginas 12 e 33.
- DEKKERS, J. C. M. Commercial application of marker and gene-assisted selection in livestock: Strategies and lessons 1,2. *Journal of Animal Science*, v. 82, n. suppl\_13, p. E313–E328, 2004. ISSN 0021-8812. Citado na página 30.
- DEROME, N. et al. A brief history of bioinformatics. *Briefings in Bioinformatics*, 2018. ISSN 1477-4054. Citado 2 vezes nas páginas 13 e 34.
- DU, Z.-Q. et al. A gene-based snp linkage map for pacific white shrimp, *litopenaeus vannamei*. *Animal genetics*, v. 41, p. 286–94, 2009. ISSN 0268-9146. Citado na página 18.
- FENG, J. Q. et al. Loss of *dmp1* causes rickets and osteomalacia and identifies a role for osteocytes in mineral metabolism. *Nat Genet*, v. 38, n. 11, p. 1310–1315, 2006. ISSN 1061-4036. Citado na página 20.
- FENG, S. et al. *Insl3/rxfp2* signaling in testicular descent. *Ann N Y Acad Sci*, v. 1160, p. 197–204, 2009. ISSN 1749-6632. Citado na página 20.

- FERLIN, A. et al. Mutations in the insulin-like factor 3 receptor are associated with osteoporosis. *J Bone Miner Res*, Amer Soc Bone & Mineral Res., v. 23, n. 5, p. 683–693, 2008. ISSN 1523-4681. Citado na página 20.
- GIBSON, J.; BISHOP, S. Use of molecular markers to enhance resistance of livestock to disease: A global approach. *Revue scientifique et technique (International Office of Epizootics)*, v. 24, p. 343–53, 05 2005. ISSN 0253-1933. Citado na página 16.
- GIRÃO, E. S.; MEDEIROS, L. P.; GIRÃO, R. N. Ocorrência e distribuição estacional de helmintos gastrintestinais de caprinos no município de teresina, piauí. *Ciência Rural*, v. 22, p. 197–202, 08 1992. ISSN 0103-8478. Citado na página 14.
- GRAVES, J. A. M. Background and Overview of Comparative Genomics. *ILAR Journal*, v. 39, n. 2-3, p. 48–65, 1998. ISSN 1084-2020. Citado na página 24.
- GROBET, L. et al. A deletion in the bovine myostatin gene causes the double-muscléd phenotype in cattle. *Nature Genetics*, v. 17, n. 1, p. 71–74, 1997. ISSN 1546-1718. Citado na página 30.
- HAEGEMAN, A. et al. Mapping and snp analysis of bovine candidate genes for meat and carcass quality. *Animal Genetics*, v. 34, n. 5, p. 349–353, 2003. ISSN 1365-2052. Citado na página 30.
- HAGEN, J. B. The origins of bioinformatics. *Nature reviews. Genetics*, v. 1, p. 231–6, 2001. ISSN 1471-0064. Citado 2 vezes nas páginas 13 e 34.
- HALL, A. E.; FIEBIG, A.; PREUSS, D. Beyond the arabidopsis genome: opportunities for comparative genomics. *Plant Physiol*, v. 129, n. 4, p. 1439–1447, 2002. ISSN 0032-0889. Citado na página 23.
- HARDISON, R. C. Comparative genomics. *PLoS Biol*, Public Library of Science, v. 1, n. 2, p. E58–E58, 2003. ISSN 1545-7885. Citado na página 23.
- HIRSCHHORN, J. N.; DALY, M. Genome-wide association studies for common diseases and complex traits. *Nature reviews. Genetics*, v. 6, p. 95–108, 2005. ISSN 1471-0064. Citado na página 19.
- Instituto Brasileiro de Geografia e Estatística. *Censo Agropecuário 2017*. 2018. <[https://biblioteca.ibge.gov.br/visualizacao/periodicos/3093/agro\\_2017\\_resultados\\_preliminares.pdf](https://biblioteca.ibge.gov.br/visualizacao/periodicos/3093/agro_2017_resultados_preliminares.pdf)>. Online; acessado em 10 Dezembro 2018. Citado 3 vezes nas páginas 12, 14 e 33.
- JOHNSTON, S. E. et al. Genome-wide association mapping identifies the genetic basis of discrete and quantitative variation in sexual weaponry in a wild sheep population. *Molecular Ecology*, v. 20, n. 12, p. 2555–2566, 2011. ISSN 1365-294X. Citado na página 20.
- KAPLAN, R. M. Drug resistance in nematodes of veterinary importance: A status report. *Trends in parasitology*, v. 20, p. 477–81, 11 2004. ISSN 1471-4922. Citado na página 14.
- KEMPER, K. E. et al. The distribution of snp marker effects for faecal worm egg count in sheep, and the feasibility of using these markers to predict genetic merit for resistance to worm infections. *Genetics Research*, Cambridge University Press, v. 93, n. 3, p. 203–219, 2011. ISSN 1469-5073. Citado na página 16.

- KLIEBENSTEIN, D. J.; GERSHENZON, J.; MITCHELL-OLDS, T. Comparative quantitative trait loci mapping of aliphatic, indolic and benzylic glucosinolate production in *Arabidopsis thaliana* leaves and seeds. *Genetics*, v. 159, n. 1, p. 359–370, 2001. ISSN 0016-6731. Citado na página 23.
- KLOOSTERMAN, A.; PARMENTIER, H.; PLOEGER, H. Breeding cattle and sheep for resistance to gastrointestinal nematodes. *Parasitology Today*, v. 8, n. 10, p. 330 – 335, 1992. ISSN 0169-4758. Citado na página 16.
- KOSHIDA, R. et al. A novel nonsense mutation in the *dmp1* gene in a Japanese family with autosomal recessive hypophosphatemic rickets. *Journal of bone and mineral metabolism*, v. 28, p. 585–90, 2010. ISSN 1435-5604. Citado na página 20.
- LESNICK, T. G. et al. A genomic pathway approach to a complex disease: Axon guidance and Parkinson disease. *PLOS Genetics*, Public Library of Science, v. 3, n. 6, p. 1–12, 2007. ISSN 1553-7404. Citado na página 19.
- LIU, S.-F.; JIANG, Y.-L.; DU, L.-X. Studies of *bmpr-ib* and *bmp15* as candidate genes for fecundity in little tailed Han sheep. *Yi chuan xue bao = Acta Genetica Sinica*, v. 30, p. 755–60, 2003. ISSN 0379-4172. Citado na página 30.
- MCMANUS, C. et al. Genetic factors of sheep affecting gastrointestinal parasite infections in the Distrito Federal, Brazil. *Veterinary Parasitology*, v. 166, n. 3, p. 308 – 313, 2009. ISSN 0304-4017. Citado na página 14.
- MELO, A. C. F. L. et al. Nematódeos resistentes a anti-helmíntico em rebanhos de ovinos e caprinos do Estado do Ceará, Brasil. *Cienc. Rural*, v. 33, n. 2, p. 339 – 344, 2003. ISSN 0103-8478. Citado 2 vezes nas páginas 12 e 33.
- MURPHY, W. J.; STANYON, R.; O'BRIEN, S. J. Evolution of mammalian genome organization inferred from comparative gene mapping. *Genome Biology*, v. 2, n. 6, p. reviews0005.1, 2001. ISSN 1474-760X. Citado na página 23.
- NCBI. *Manual do usuário de aplicativos de linha de comando do BLAST® [Internet]*. [S.l.], 2008. Disponível em: <<https://www.ncbi.nlm.nih.gov/books/NBK279670/>>. Acesso em: 15 jan. 2019. Citado na página 26.
- NOVA, L. E. V. et al. Resistência de nematódeos aos anti-helmínticos nitroxinil 34% e ivermectina 1% em rebanho ovino no município de São João do Itaipava, Paraná. *Revista Brasileira de Higiene e Sanidade Animal*, v. 8, n. 1, p. 160–171, 2014. ISSN 1981 –2965. Citado 2 vezes nas páginas 12 e 33.
- PATERSON, A. H. et al. Comparative genomics of plant chromosomes. *Plant Cell*, American Society of Plant Physiologists, v. 12, n. 9, p. 1523–1540, 2000. ISSN 1040-4651. Citado na página 23.
- PAVY, N. et al. Automated SNP detection from a large collection of white spruce expressed sequences: contributing factors and approaches for the categorization of SNPs. *BMC Genomics*, v. 7, n. 1, p. 174, 2006. ISSN 1471-2164. Citado na página 18.
- PERTSEMLIDIS, A.; FONDON, J. W. R. Having a blast with bioinformatics (and avoiding blastphemy). *Genome Biol*, BioMed Central, v. 2, n. 10, p. REVIEWS2002–REVIEWS2002, 2001. ISSN 1474-760X. Citado na página 27.

- P.F.BRYNE; M.D.MCMULLEN. Defining genes for agricultural traits: Qtl analysis and the candidate gene approach. *Probe*, v. 7, p. 24–27, 1996. Citado na página 30.
- PIERCE, B. A. *Genética: um enfoque conceitual*. 3a edição. ed. [S.l.]: GUANABARA KOOGAN, 2011. ISBN 9788527716642. Citado 2 vezes nas páginas 21 e 22.
- PLENGE, R. M. et al. Traf1-c5 as a risk locus for rheumatoid arthritis — a genomewide study. *New England Journal of Medicine*, v. 357, n. 12, p. 1199–1209, 2007. ISSN 1533-4406. Citado na página 19.
- QUINNELL, R. J. Genetics of susceptibility to human helminth infection. *International Journal for Parasitology*, v. 33, n. 11, p. 1219 – 1231, 2003. ISSN 0020-7519. Molecular and Cellular Biology of Helminth Parasites. Citado na página 16.
- RAFALSKI, J. Applications of single nucleotide polymorphisms in crop genetics. *Current opinion in plant biology*, v. 5, p. 94–100, 05 2002. ISSN 1369-5266. Citado na página 18.
- RAMOS, C. I. et al. Epidemiologia das helmintoses gastrintestinais de ovinos no planalto catarinense. *Cienc. Rural*, v. 34, p. 1889 – 1895, 2004. ISSN 0103-8478. Citado na página 14.
- SACHIDANANDAM, R. et al. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature*, v. 409, p. 928–933, 01 2001. ISSN 1476-4687. Citado na página 18.
- SADDIQI, H. A. et al. Small ruminant resistance against gastrointestinal nematodes: A case of haemonchus contortus. *Parasitology research*, v. 109, p. 1483–500, 2011. ISSN 1432-1955. Citado na página 15.
- SANKOFF, D.; NADEAU, J. *Comparative Genomics: Empirical and Analytical Approaches to Gene Order Dynamics, Map Alignment and the Evolution of Gene Families*. 1a edição. ed. [S.l.]: Kluwer Academic, 2000. ISBN 0792365836. Citado na página 23.
- SANTIAGO, M. A. M.; BENEVENGA, S. F.; COSTA, U. C. Epidemiologia e controle da helmintose ovina no município de Itaquí, Rio Grande do Sul. *Pesq. Agropec. Bras., Ser. Vet.*, v. 11, n. 9, p. 1–7, 1976. ISSN 1678-3991. Citado na página 14.
- SAXENA, R. et al. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science (New York, N.Y.)*, v. 316, p. 1331–6, 2007. ISSN 2375-723X. Citado na página 19.
- SAYERS, G.; SWEENEY, T. Gastrointestinal nematode infection in sheep – a review of the alternatives to anthelmintics in parasite control. *Animal Health Research Reviews*, Cambridge University Press, v. 6, n. 2, p. 159–171, 2005. ISSN 1475-2654. Citado na página 14.
- SCHMIDT, R. H. Plant genome evolution: lessons from comparative genomics at the dna level. *Plant Molecular Biology*, v. 48, p. 21–37, 2002. ISSN 1573-5028. Citado na página 23.
- SRETER, T.; KASSAI, T.; TAKACS, E. The heritability and specificity of responsiveness to infection with haemonchus contortus in sheep. *International Journal for Parasitology*, v. 24, n. 6, p. 871 – 876, 1994. ISSN 0020-7519. Citado na página 15.

- STEAR, M.; MURRAY, M. Genetic resistance to parasitic disease: particularly of resistance in ruminants to gastrointestinal nematodes. *Veterinary Parasitology*, v. 54, n. 1, p. 161 – 176, 1994. ISSN 0304-4017. Special Issue: 14th W.A.A.V.P. Conference. Citado 2 vezes nas páginas 15 e 16.
- STEPHENS, J. Haplotype variation and linkage disequilibrium in 313 human genes (vol 293, pg 489, 2001). *Science*, v. 293, p. 1048–1048, 08 2001. ISSN 1095-9203. Citado na página 18.
- TORRES-ACOSTA, J.; HOSTE, H. Alternative or improved methods to limit gastro-intestinal parasitism in grazing sheep and goats. *Small Ruminant Research*, v. 77, n. 2, p. 159 – 173, 2008. ISSN 0921-4488. Special Issue: Sheep and Goat Farming: grazing systems of production and development. Citado 2 vezes nas páginas 12 e 33.
- VIGNAL, A. et al. A review on snp and other types of molecular markers and their use in animal genetics. *Genetics Selection Evolution*, v. 34, n. 3, p. 275, 2002. ISSN 1297-9686. Citado 2 vezes nas páginas 16 e 17.
- VRIES, A. G. de et al. The role of major genes and dna technology in selection for meat quality in pigs. *Meat Science*, v. 49, p. S245–S255, 1998. ISSN 0309-1740. Citado na página 30.
- WALLER, P. et al. The prevalence of anthelmintic resistance in nematode parasites of sheep in southern latin america: General overview. *Veterinary Parasitology*, v. 62, n. 3, p. 181 – 187, 1996. ISSN 0304-4017. Citado 2 vezes nas páginas 12 e 33.
- WALLER, P. J. Anthelmintic resistance. *Veterinary Parasitology*, v. 72, n. 3, p. 391 – 412, 1997. ISSN 0304-4017. Citado na página 14.
- WAN, Y. et al. Secondary cell wall composition and candidate gene expression in developing willow (*salix purpurea*) stems. *Planta*, v. 239, n. 5, p. 1041–1053, 2014. ISSN 1432-2048. Citado na página 37.
- WEI, L. et al. Comparative genomics approaches to study organism similarities and differences. *Journal of Biomedical Informatics*, v. 35, n. 2, p. 142 – 150, 2002. ISSN 1532-0464. Citado na página 23.
- YANG, W. et al. Review on the development of genotyping methods for assessing farm animal diversity. *Journal of animal science and biotechnology*, v. 4, p. 2, 2013. ISSN 2049-1891. Citado na página 17.
- YUAN, F. P. et al. The role of rxfp2 in mediating androgen-induced inguinoscrotal testis descent in lh receptor knockout mice. *Reproduction*, v. 139, n. 4, p. 759–769, 2010. ISSN 1741-7899. Citado na página 20.
- ZHANG, H. et al. Progress of genome wide association study in domestic animals. *Journal of Animal Science and Biotechnology*, v. 3, n. 1, p. 26, 2012. ISSN 2049-1891. Citado na página 19.
- ZHANG, J. et al. Locating genomic regions associated with components of drought resistance in rice: Comparative mapping within and across species. *Theoretical and Applied Genetics*, v. 103, p. 19–29, 2001. ISSN 1432-2242. Citado na página 23.

ZHAO, X. et al. A novel nonsense mutation in the *dmp1* gene identified by a genome-wide association study is responsible for inherited rickets in Corriedale sheep. *PLoS One*, Public Library of Science, v. 6, n. 7, p. e21739–e21739, 2011. ISSN 1932-6203. Citado na página 20.

### 3.5 Considerações finais

O uso da informação de SNPs associados com características de interesse econômico aliados com técnicas de comparação de genomas pode respaldar a descoberta de genes importantes para o melhoramento genético animal. A metodologia aqui usada pode ser expandida para diferentes espécies além do *Ovis aries*, onde trabalhos aprofundados sobre os genes candidatos poderão contribuir na melhora do desempenho dos animais e na qualidade dos produtos oriundos dos mesmos. Para tanto, é necessário encontrar mais SNPs associados a fenótipos de interesse para que se possam aumentar o número de informações sobre genes candidatos e aprofundar os estudos para encontrar os principais genes que comandam a expressão e controle das características.

# Apêndices



# APÊNDICE A – Genes Encontrados nas Janelas do Ovis Aries

Neste apêndice, estão listados todos os genes encontrados nas janelas da espécie *Ovis aries*, localização e classificação de acordo com envolvimento ou não em processos imunológicos.

Gene	Localização	Processo Imunológico
ENSOARG00000017817	Chr21 46,557,298-46,656,557	Sem Relação
ENSOARG00000018468	Chr21 46,672,617-46,674,854	<b>Com Relação</b>
ENSOARG00000018678	Chr21 46,721,228-46,797,750	Sem Relação
ENSOARG00000000456	Chr21 46,818,359-46,839,647	Sem Relação
ENSOARG00000017611	Chr21 46,325,179-46,329,754	Sem Relação
ENSOARG00000000839	Chr21 46,860,749-46,883,416	Sem Relação
ENSOARG00000000953	Chr21 47,016,031-47,019,713	Sem Relação
ENSOARG00000001052	Chr21 47,146,103-47,308,294	Sem Relação
ENSOARG00000001571	Chr21 47,597,876-47,626,025	Sem Relação
ENSOARG00000015703	Chr21 47,709,314-47,709,496	Sem Relação
ENSOARG00000001778	Chr21 47,746,255-47,802,320	Sem Relação
ENSOARG00000002287	Chr21 47,862,666-47,898,078	Sem Relação
ENSOARG00000001965	Chr21 47,814,507-47,833,642	Sem Relação
ENSOARG00000001324	Chr21 47,528,820-47,556,068	Sem Relação
ENSOARG00000001371	Chr21 47,576,076-47,584,581	Sem Relação
ENSOARG000000015696	Chr21 47,631,203-47,632,150	<b>Com Relação</b>
ENSOARG00000002554	Chr21 47,911,171-47,935,403	<b>Com Relação</b>
ENSOARG00000002794	Chr21 48,179,420-48,271,476	Sem Relação

Tabela 7: Genes encontrados na janela do SNP: OAR21\_52583090.1

Gene	Localização	Processo Imunológico
ENSOARG00000010120	Chr12 13,398,010-13,706,664	Sem Relação

Tabela 8: Genes encontrados na janela do SNP: OAR12\_17005059.1

Gene	Localização	Processo Imunológico
ENSOARG00000004773	Chr8 83,449,093-83,583,173	Sem Relação
ENSOARG00000004781	Chr8 83,627,065-83,633,679	<b>Com Relação</b>
ENSOARG00000004788	Chr8 83,818,709-84,599,387	<b>Com Relação</b>
ENSOARG00000004794	Chr8 84,838,731-84,906,013	Sem Relação
ENSOARG00000004799	Chr8 85,118,793-85,143,936	Sem Relação
ENSOARG00000002798	Chr8 84,889,785-84,890,160	Sem Relação

Tabela 9: Genes encontrados na janela do SNP: OAR8\_91253490\_X.1.



Gene	Localização	Processo Imunológico
ENSOARG00000008608	Chr2 30,912,297-31,124,598	Sem Relação
ENSOARG00000008650	Chr2 31,653,188-31,684,127	Sem Relação
ENSOARG00000008668	Chr2 31,709,019-31,765,595	Sem Relação
ENSOARG00000008634	Chr2 31,239,795-31,476,634	<b>Com Relação</b>
ENSOARG00000008621	Chr2 31,141,125-31,163,045	<b>Com Relação</b>
ENSOARG00000008683	Chr2 31,785,406-31,790,389	<b>Com Relação</b>
ENSOARG00000008706	Chr2 31,865,552-31,962,542	Sem Relação

Tabela 10: Genes encontrados na janela do SNP: s23985.1.

Gene	Localização	Processo Imunológico
ENSOARG00000004787	Chr6 39,939,556-40,091,549	<b>Com Relação</b>
ENSOARG00000004640	Chr6 39,766,345-39,768,351	Sem Relação
ENSOARG00000005149	Chr6 40,156,628-40,399,467	Sem Relação

Tabela 11: Genes encontrados na janela do SNP: s26633.1.

Gene	Localização	Processo Imunológico
ENSOARG00000002884	Chr2 138,639,228-138,648,094	Sem Relação
ENSOARG00000003284	Chr2 138,805,177-138,838,240	Sem Relação
ENSOARG00000003592	Chr2 138,990,254-139,154,704	Sem Relação
ENSOARG00000004000	Chr2 139,245,027-139,331,632	Sem Relação
ENSOARG00000004104	Chr2 139,371,222-139,382,831	Sem Relação
ENSOARG00000003082	Chr2 138,727,404-138,753,440	Sem Relação
ENSOARG00000002804	Chr2 138,424,294-138,625,167	<b>Com Relação</b>
ENSOARG000000019650	Chr2 138,721,646-138,722,371	Sem Relação
ENSOARG00000003388	Chr2 138,842,626-138,856,360	<b>Com Relação</b>
ENSOARG00000003872	Chr2 139,186,849-139,193,920	Sem Relação
ENSOARG00000004085	Chr2 139,349,795-139,356,438	Sem Relação
ENSOARG00000004285	Chr2 139,486,616-139,831,507	Sem Relação
ENSOARG00000002906	Chr2 138,648,159-138,653,917	Sem Relação
ENSOARG00000003174	Chr2 138,758,565-138,779,078	Sem Relação
ENSOARG00000003402	Chr2 138,859,320-138,879,396	Sem Relação
ENSOARG00000004206	Chr2 139,391,357-139,458,885	Sem Relação
ENSOARG00000003017	Chr2 138,681,642-138,697,123	Sem Relação
ENSOARG000000019653	Chr2 140,017,409-140,017,636	Sem Relação
ENSOARG00000004301	Chr2 140,068,741-140,315,412	<b>Com Relação</b>
ENSOARG00000004285	Chr2 139,486,616-139,831,507	Sem Relação
ENSOARG000000019658	Chr2 140,432,670-140,433,650	Sem Relação

Tabela 12: Genes encontrados na janela do SNP: OAR2\_148350187.1.

## APÊNDICE B – Alinhamentos que ocorreram entre genes

Neste apêndice, estão listados todos o genes que alinharam entre a espécie *Ovis aires* com as espécies *Sus scrofa*, *Capra hircus*, *Bos taurus*, *Mus musculos* e *Homo sapiens*.

Tabela 13: Genes com relação com processo imunológico-OAR21\_52583090.1

Gene	Localização	Espécie
ENSSSCG00000039699	Chr2 2,328,524-2,332,633	Sus scrofa

Tabela 14: Genes sem relação com processo imunológico-OAR21\_52583090.1

Gene	Localização	Espécie
ENSCHIG0000000879	Chr1 139,040,362-139,041,375	Capra hircus
ENSCHIG0000001410	Chr13 39,628,318-39,630,821	Capra hircus
ENSCHIG0000003238	Chr2 125,903,097-125,904,217	Capra hircus
ENSCHIG0000008770	Chr4 103,398,988-103,401,571	Capra hircus
ENSG00000239900	Chr22 40,346,500-40,390,463	Homo sapiens
ENSSSCG00000032624	Chr5 83,354,488-83,356,983	Sus scrofa
ENSSSCG00000038939	Chr14 9,699,489-9,702,689	Sus scrofa
ENSSSCG00000037695	Chr7 45,298,705-45,299,811	Sus scrofa
ENSSSCG00000034665	Chr13 38,542,781-38,544,856	Sus scrofa
ENSSSCG00000039035	Chr1 143,696,607-143,699,105	Sus scrofa
ENSSSCG00000035547	Chr15 16,156,195-16,158,044	Sus scrofa
ENSSSCG00000040911	Chr6 108,635,250-108,639,038	Sus scrofa
ENSSSCG00000037694	Chr3 80,062,731-80,063,396	Sus scrofa
ENSSSCG00000040141	Chr2 40,990,843-40,994,349	Sus scrofa
ENSSSCG00000033271	Chr9 51,337,178-51,338,098	Sus scrofa
ENSSSCG00000039209	Chr15 28,834,319-28,837,216	Sus scrofa
ENSSSCG00000037576	Chr11 1,443,136-1,445,715	Sus scrofa
ENSSSCG00000034985	Chr16 66,417,576-66,419,750	Sus scrofa
ENSSSCG00000032975	Chr14 82,099,487-82,100,569	Sus scrofa
ENSSSCG00000037444	Chr16 28,649,720-28,653,045	Sus scrofa
ENSSSCG00000034223	Chr3 31,466,874-31,470,623	Sus scrofa
ENSSSCG00000031466	Chr16 50,186,591-50,189,123	Sus scrofa
ENSSSCG00000035590	Chr13 154,546,664-154,547,941	Sus scrofa
ENSSSCG00000033645	Chr4 93,177,170-93,179,068	Sus scrofa
ENSSSCG00000037679	Chr7 94,157,732-94,163,433	Sus scrofa
ENSSSCG00000035697	Chr4 114,936,589-114,940,929	Sus scrofa
ENSSSCG00000039159	Chr13 162,505,609-162,508,349	Sus scrofa
ENSSSCG00000033884	Chr13 162,720,267-162,721,322	Sus scrofa
ENSSSCG00000039422	Chr7 63,394,845-63,397,710	Sus scrofa
ENSSSCG00000034560	ChrX 41,369,159-41,379,383	Sus scrofa
ENSSSCG00000034659	Chr3 30,480,255-30,481,509	Sus scrofa
ENSSSCG00000035060	Chr8 103,028,687-103,032,398	Sus scrofa
ENSSSCG00000032736	ChrX 41,200,707-41,731,258	Sus scrofa
ENSSSCG00000033364	Chr2 107,753,052-107,754,077	Sus scrofa
ENSSSCG00000037664	Chr6 82,672,445-82,674,046	Sus scrofa
ENSSSCG00000035961	Chr13 117,363,066-117,363,988	Sus scrofa
ENSSSCG00000036671	Chr1 39,027,473-39,578,423	Sus scrofa
ENSSSCG00000033555	Chr7 113,797,561-113,800,875	Sus scrofa
ENSSSCG00000036227	Chr13 29,479,280-29,484,683	Sus scrofa
ENSSSCG00000031243	ChrX 41,032,365-41,376,796	Sus scrofa
ENSSSCG00000031243	ChrX 41,032,365-41,376,796	Sus scrofa
ENSMUSG00000022407	Chr15 80,948,490-80,970,946	Mus musculus

Tabela 15: Genes com relação com processo imunológico-OAR12\_17005059.1

Gene	Localização	Espécie
ENSMUSG00000032518	Chr9 120,127,689-120,132,369	Mus musculus

Tabela 16: Genes sem relação com processo imunológico-OAR12\_17005059.1

Gene	Localização	Espécie
ENSG00000168028	Chr3 39,406,689-39,421,542	Homo sapiens
ENSBTAG00000009757	Chr22 12,684,159-12,710,177	Bos taurus
ENSCHIG00000007093	Chr7 1,943,664-1,944,992	Capra hircus
ENSCHIG00000011924	Chr22 12,541,215-12,552,935	Capra hircus
ENSSSCG00000037694	Chr3 80,062,731-80,063,396	Sus scrofa
ENSSSCG00000033271	Chr9 51,337,178-51,338,098	Sus scrofa
ENSSSCG00000034223	Chr3 31,466,874-31,470,623	Sus scrofa
ENSSSCG00000033645	Chr4 93,177,170-93,179,068	Sus scrofa
ENSSSCG00000032736	ChrX 41,200,707-41,731,258	Sus scrofa
ENSSSCG00000035961	Chr13 117,363,066-117,363,988	Sus scrofa
ENSSSCG00000037236	Chr13 24,072,510-24,083,686	Sus scrofa
ENSSSCG00000031243	ChrX 41,032,365-41,376,796	Sus scrofa
ENSMUSG00000096808	Chr16 15,317,458-15,318,957	Mus musculus
ENSMUSG00000094472	Chr16 15,317,458-15,318,957	Mus musculus
ENSMUSG00000096463	Chr14 73,661,225-73,662,721	Mus musculus
ENSMUSG00000094030	Chr16 82,828,382-82,829,881	Mus musculus

Tabela 17: Genes sem relação com processo imunológico-OAR8\_91253490\_X.1

Gene	Localização	Espécie
ENSBTAG00000014226	Chr6 16,671,257-16,675,901	Bos taurus
ENSMUSG00000062006	Chr3 130,726,831-130,730,398	Mus musculus
ENSG00000109475	Chr4 108,620,566-108,630,412	Homo sapiens
ENSSSCG00000037576	Chr11 1,443,136-1,445,715	Sus scrofa
ENSSSCG00000034223	Chr3 31,466,874-31,470,623	Sus scrofa
ENSSSCG00000035590	Chr13 154,546,664-154,547,941	Sus scrofa
ENSSSCG00000033645	Chr4 93,177,170-93,179,068	Sus scrofa
ENSSSCG00000037679	Chr7 94,157,732-94,163,433	Sus scrofa
ENSSSCG00000039159	Chr13 162,505,609-162,508,349	Sus scrofa
ENSSSCG00000009146	Chr8 113,467,779-113,471,982	Sus scrofa
ENSSSCG00000034659	Chr3 30,480,255-30,481,509	Sus scrofa
ENSSSCG00000032736	ChrX 41,200,707-41,731,258	Sus scrofa
ENSSSCG00000036671	Chr1 39,027,473-39,578,423	Sus scrofa
ENSSSCG00000036227	Chr13 29,479,280-29,484,683	Sus scrofa
ENSSSCG00000031243	ChrX 41,032,365-41,376,796	Sus scrofa
ENSCHIG00000001904	Chr20 64,534,765-64,535,118	Capra hircus
ENSCHIG00000003828	Chr25 22,183,109-22,183,453	Capra hircus
ENSCHIG00000007093	Chr7 1,943,664-1,944,992	Capra hircus
ENSCHIG00000013243	Chr6 16,990,826-16,994,418	Capra hircus

Tabela 18: Genes sem relação com processo imunológico- OAR2\_148350187.1

Gene	Localização	Espécie
ENSCHIG00000003189	Chr18 3,759,282-3,759,506	Capra hircus
ENSCHIG00000003204	Chr2 109,681,111-109,681,836	Capra hircus
ENSCHIG00000004959	Chr8 109,699,922-109,702,525	Capra hircus
ENSCHIG000000015105	Chr2 107,977,283-107,978,262	Capra hircus
ENSCHIG000000016789	Chr14 17,758,452-17,759,625	Capra hircus
ENSMUSG000000094472	Chr16 15,317,458-15,318,957	Mus musculus
ENSMUSG000000096463	Chr14 73,661,225-73,662,721	Mus musculus
ENSMUSG000000094030	Chr16 82,828,382-82,829,881	Mus musculus
ENSMUSG000000034780	Chr2 67,565,871-68,122,689	Mus musculus
ENSMUSG000000027088	Chr2 69,789,623-69,800,005	Mus musculus
ENSSSCG000000037695	Chr7 45,298,705-45,299,811	Sus scrofa
ENSSSCG000000035547	Chr15 16,156,195-16,158,044	Sus scrofa
ENSSSCG000000015936	Chr15 76,007,208-76,014,008	Sus scrofa
ENSSSCG000000037576	Chr11 1,443,136-1,445,715	Sus scrofa
ENSSSCG000000034223	Chr3 31,466,874-31,470,623	Sus scrofa
ENSSSCG000000006546	Chr4 95,231,252-95,284,502	Sus scrofa
ENSSSCG000000034351	Chr15 74,236,415-74,340,060	Sus scrofa
ENSSSCG000000033645	Chr4 93,177,170-93,179,068	Sus scrofa
ENSSSCG000000039159	Chr13 162,505,609-162,508,349	Sus scrofa
ENSSSCG000000033884	Chr13 162,720,267-162,721,322	Sus scrofa
ENSSSCG000000032736	ChrX 41,200,707-41,731,258	Sus scrofa
ENSSSCG000000035961	Chr13 117,363,066-117,363,988	Sus scrofa
ENSSSCG000000036671	Chr1 39,027,473-39,578,423	Sus scrofa
ENSSSCG000000036227	Chr13 29,479,280-29,484,683	Sus scrofa
ENSSSCG000000031243	ChrX 41,032,365-41,376,796	Sus scrofa
ENSG000000144362	Chr2 169,694,454-169,701,708	Homo sapiens
ENSG000000172318	Chr2 167,868,948-167,874,041	Homo sapiens
ENSBTAG00000001092	Chr2 26,613,246-26,618,722	Bos taurus

Tabela 19: Genes sem relação com processo imunológico- OAR2\_148350187.1

Gene	Localização	Espécie
ENSCHIG00000000879	Chr1 139,040,362-139,041,375	Capra hircus
ENSCHIG00000003238	Chr2 125,903,097-125,904,217	Capra hircus
ENSCHIG00000004959	Chr8 109,699,922-109,702,525	Capra hircus
ENSCHIG00000004959	Chr8 109,699,922-109,702,525	Capra hircus
ENSMUSG00000096463	Chr14 73,661,225-73,662,721	Mus musculus
ENSMUSG00000094030	Chr16 82,828,382-82,829,881	Mus musculus
ENSSSCG00000037695	Chr7 45,298,705-45,299,811	Sus scrofa
ENSSSCG00000037694	Chr3 80,062,731-80,063,396	Sus scrofa
ENSSSCG00000035547	Chr15 16,156,195-16,158,044	Sus scrofa
ENSSSCG00000037694	Chr3 80,062,731-80,063,396	Sus scrofa
ENSSSCG00000037576	Chr11 1,443,136-1,445,715	Sus scrofa
ENSSSCG00000032975	Chr14 82,099,487-82,100,569	Sus scrofa
ENSSSCG00000034223	Chr3 31,466,874-31,470,623	Sus scrofa
ENSSSCG00000033645	Chr4 93,177,170-93,179,068	Sus scrofa
ENSSSCG00000037679	Chr7 94,157,732-94,163,433	Sus scrofa
ENSSSCG00000033884	Chr13 162,720,267-162,721,322	Sus scrofa
ENSSSCG00000032736	ChrX 41,200,707-41,731,258	Sus scrofa
ENSSSCG00000033364	Chr2 107,753,052-107,754,077	Sus scrofa
ENSSSCG00000035961	Chr13 117,363,066-117,363,988	Sus scrofa
ENSSSCG00000036671	Chr1 39,027,473-39,578,423	Sus scrofa
ENSSSCG00000035080	Chr12 44,957,152-44,962,723	Sus scrofa
ENSSSCG00000031243	ChrX 41,032,365-41,376,796	Sus scrofa
ENSBTAG00000023106	Chr1 139,422,600-139,432,042	Bos taurus
ENSBTAG00000022498	Chr4 43,148,370-43,148,895	Bos taurus
ENSBTAG00000032432	Chr14 66,645,755-66,646,231	Bos taurus
ENSBTAG00000032775	Chr11 21,528,543-21,529,010	Bos taurus

Tabela 20: Genes com relação com processo imunológico - s23985

Gene	Localização	Espécie
ENSBTAG00000014534	Chr9 13,011,231-13,020,524	Bos taurus
Eef1a1 ENSMUSG00000037742	Chr9 78,478,449-78,489,151	Mus musculus
EEF1A1 ENSG00000156508	Chr6 73,515,750-73,523,797	Homo sapiens

Tabela 21: Genes sem relação com processo imunológico - s23985

Gene	Localização	Espécie
ENSBTAG00000035926	12 783,000-784,421	Bos taurus
ENSSSCG00000037694	Chr3 80,062,731-80,063,396	Sus scrofa
ENSSSCG00000037576	Chr11 1,443,136-1,445,715	
ENSSSCG00000038035	Chr6 81,509,567-81,519,072	Sus scrofa
ENSSSCG00000034223	Chr3 31,466,874-31,470,623	Sus scrofa
ENSSSCG00000034113	Chr9 25,500,470-25,503,075	Sus scrofa
ENSSSCG00000004489	Chr1 92,421,230-92,472,785	Sus scrofa
ENSSSCG00000033645	Chr4 93,177,170-93,179,068	Sus scrofa
ENSSSCG00000037679	Chr7 94,157,732-94,163,433	Sus scrofa
ENSSSCG00000039159	Chr13 162,505,609-162,508,349	Sus scrofa
ENSSSCG00000033884	Chr13 162,720,267-162,721,322	Sus scrofa
ENSSSCG00000031239	Chr15 24,707,846-24,709,454	Sus scrofa
ENSSSCG00000034659	Chr3 30,480,255-30,481,509	Sus scrofa
ENSSSCG00000032736	ChrX 41,200,707-41,731,258	Sus scrofa
ENSSSCG00000040401	Chr17 62,515,417-62,526,190	Sus scrofa
ENSSSCG00000036671	Chr1 39,027,473-39,578,423	Sus scrofa
ENSSSCG00000036227	Chr13 29,479,280-29,484,683	Sus scrofa
ENSSSCG00000031243	ChrX 41,032,365-41,376,796	Sus scrofa
ENSSSCG00000031243	ChrX 41,032,365-41,376,796	Sus scrofa
ENSSSCG00000031243	ChrX 41,032,365-41,376,796	Sus scrofa
ENSMUSG00000094472	Chr16 15,317,458-15,318,957	Mus musculus
ENSMUSG00000096463	Chr14 73,661,225-73,662,721	Mus musculus
ENSMUSG00000094030	Chr16 82,828,382-82,829,881	Mus musculus
ENSMUSG00000016349	Chr2 181,147,653-181,157,014	Mus musculus
ENSG00000148120	Chr9 94,726,701-95,087,218	Homo sapiens
ENSG00000175826	Chr17 7,243,591-7,252,491	Homo sapiens
ENSG00000101210	Chr20 63,488,012-63,499,322	Homo sapiens
ENSCHIG00000004959	Chr8 109,699,922-109,702,525	Capra hircus
ENSCHIG00000007885	Chr1 37,867,171-37,868,969	Capra hircus
ENSCHIG00000008770	Chr4 103,398,988-103,401,571	Capra hircus
ENSCHIG00000010087	Chr9 183,177-185,304	Capra hircus
ENSCHIG00000016589	Chr13 53,475,234-53,484,361	Capra hircus
ENSCHIG00000021742	Chr8 10,835,516-10,836,282	Capra hircus