UNIVERSIDADE DE SÃO PAULO
FACULDADE DE ZOOTECNIA E ENGENHARIA DE ALIMENTOS

YURI PEREIRA EFREM NATIVIDADE

Seleção Genômica Multirracial em Bovinos de Corte

Pirassununga

2017

YURI PEREIRA EFREM NATIVIDADE

Seleção Genômica Multirracial em Bovinos de Corte

Dissertação apresentada à Faculdade de Zootecnia e Engenharia de Alimentos da Universidade de São Paulo, como parte dos requisitos para a obtenção do Título de Mestre em Ciências.

Área de Concentração: Qualidade e Produtividade Animal

Orientador: Prof. Dr. Joanir Pereira Eler

Ficha catalográfica elaborada pelo
Serviço de Biblioteca e Informação, FZEA/USP,
com os dados fornecidos pelo(a) autor(a)

N278s Natividade, Yuri Pereira Efrem
 Seleção genômica multirracial em bovinos de corte
 / Yuri Pereira Efrem Natividade ; orientador Joanir
 Pereira Eler. -- Pirassununga, 2017.
 40 f.

 Dissertação (Mestrado - Programa de Pós-Graduação
 em Zootecnia) -- Faculdade de Zootecnia e
 Engenharia de Alimentos, Universidade de São Paulo.

 1. melhoramento genético animal. 2. seleção
 genômica. 3. populações multirraciais. 4. bovinos de
 corte. I. Eler, Joanir Pereira, orient. II. Título.

YURI PEREIRA EFREM NATIVIDADE

Seleção Genômica Multirracial em Bovinos de Corte

Dissertação apresentada à Faculdade de Zootecnia e Engenharia de Alimentos da Universidade de São Paulo, como parte dos requisitos para a obtenção do Título de Mestre em Ciências.

Área de Concentração: Qualidade e Produtividade Animal

Data de aprovação: ____/____/____

Banca Examinadora:

Nome completo do 1º Examinador – Presidente da Banca Examinadora
Prof. + titulação (caso tenha) + instituição a que pertence - Orientador(a)

Nome completo do 2º Examinador
Prof. + titulação (caso tenha) + instituição a que pertence

Nome completo do 3º Examinador
Prof. + titulação (caso tenha) + instituição a que pertence

DEDICATÓRIA

Aos meus pais, Darlene e Bosco.

À minha irmã Anna Carolina e sobrinhas Júlia e Marina.

À minha avó Vicentina e em memória dos avos João, Dalva e Jairo.

À minha namorada Dayana.

AGRADECIMENTOS

Ao meu orientador, Prof. Joanir Pereira Eler, por toda a atenção desde o primeiro contato via e-mail, os ensinamentos a mim dedicados, as lições de vida passadas no dia-a-dia de maneira despretensiosa mas não por isso menos importantes.

Ao Prof. Ricardo Vieira Ventura, pelos conhecimentos na área da genômica, pelas conversas sempre francas e fraternais, pela paciência e pela perseverança.

Aos amigos e colegas do GMAB, pela convivência diária, experiências repassadas e vividas, pelo apoio e amizade nesses anos de mestrado.

Aos outros amigos feitos e encontrados nessas terras Pirassununguenses que tornaram meus dias mais prazerosos e alegres.

RESUMO

NATIVIDADE, Y. P. E. **Seleção genômica multirracial em bovinos de corte**. 2017. 40 f. Dissertação (Mestrado) – Faculdade de Zootecnia e Engenharia de Alimentos, Universidade de São Paulo, Pirassununga, 2017.

A seleção genômica é a mais moderna tecnologia no que tange a utilização de marcadores genéticos para a seleção e melhoramento genético de animais domésticos. Em síntese a metodologia consiste em uma seleção assistida por marcadores em uma escala ampla do Genoma e foi proposta por Meuwissen et al. (2001). O desequilíbrio de ligação (DL) entre os marcadores e loci de características quantitativas (QTL) e a composição da população referência são pontos chave para a confiabilidade da seleção genômica. À medida que os indivíduos se distanciam geneticamente, o DL entre marcadores e QTL diminui, o que dificulta a aplicação da seleção genômica em populações de animais multirraciais e explica o fato de até hoje a maior parte das pesquisas se concentrarem na estimação de valores genômicos apenas para animais de raças puras. A aplicação dessa metodologia em populações multirraciais de bovinos constitui uma possibilidade real de grandes avanços no melhoramento genético do rebanho brasileiro de bovinos de corte.

Palavras-chave: Melhoramento genético animal, seleção genômica, populações multirraciais.

ABSTRACT

NATIVIDADE, Y. P. E. **Multibreed genomic selection in beef cattle**. 2017. 40 f. M. Sc. Dissertasion – Faculdade de Zootecnia e Engenharia de Alimentos, Universidade de São Paulo, Pirassununga, 2017.

Genomic Selection is the newest technology into the use of genetic markers at the animal breeding. In sinthesys the methodology consists in a marker assisted selection in a genome wide scale, it was pourpused by Mewissent et al. (2001). The linkage disequilibrium between the marker and quantitative trait loci (QTL) and the composition of the reference population are key points to the reablility of the genomic selection. Once the individuals get genetically distants, the DL between markers and QTLs decays, what turns hard the aplication of genomic selection in multirracial populations of animals and explains the fact of untill today the major part of researchs are dedicated to the estimation of genomic breeding values only to purebreed animals. The aplication of this methodology in multibreed cattle populations consists in a real possibility to reach greats advances in genetic improvement of the brazilian beef cattle heard.

Key-words: Animal breeding, genomic selection, multibreed populations.

LISTA DE FIGURAS

Figura 1 – Estrutura Populacional.....	20
Figura 2 – Estrutura e validação dos cenários de populações recentes.....	25
Figura 3 - Comparação dos <i>Manhattan Plots</i> dos diferentes cenários.....	30

LISTA DE TABELAS

Tabela 1 – Cenários de simulação, de acordo com a herdabilidade e nº de QTLs.....	17
Tabela 2 – Parâmetros utilizados para simulação das populações de Bovinos de Corte.....	22
Tabela 3 – Dados utilizados para simulação do genoma do boi doméstico, nº de marcadores SNPs e loci de efeito quantitativo por cromossoma.....	23
Tabela 4 – Desequilíbrio de ligação médio (r^2) e desvios-padrão, na última geração de cada grupamento racial, no cenário de 63 QTLs e $h^2 = 0,15$	26
Tabela 5 – Desequilíbrio de ligação médio (r^2) e desvios-padrão, na última geração de cada grupamento racial, no cenário de 63 QTLs e $h^2 = 0,35$	27
Tabela 6 – Desequilíbrio de ligação médio (r^2) e desvios-padrão, na última geração de cada grupamento racial, no cenário de 1000 QTLs e $h^2 = 0,15$	28
Tabela 7 – Desequilíbrio de ligação médio (r^2) e desvios-padrão, na última geração de cada grupamento racial, no cenário de 1000 QTLs e $h^2 = 0,30$	29
Tabela 8 – Acurácia da predição do Valor Genômico Estimado utilizando 1.000 QTLs e $h^2 = 0,30$	31
Tabela 9 – Acurácia da predição do Valor Genômico Estimado utilizando 1.000 QTLs e $h^2 = 0,15$	32
Tabela 10 – Acurácia da predição do Valor Genômico Estimado utilizando 63 QTLs e $h^2 = 0,30$	32

Tabela 11 – Acurácia da predição do Valor Genômico Estimado utilizando 63 QTLs e $h^2 =$

0,30.....33

LISTA DE ABREVIATURAS E SIMBOLOS

- SNP - Single Nucleotide Polymorphism (Nucleotídeo de polimorfismo único)
- QTL - Quantitative Trait Loci (Loci de característica quantitativa)
- DL - Desequilíbrio de Ligação
- h^2 - Coeficiente de herdabilidade
- cM - CentiMorgans
- bp - Base Pair (Par de Bases)
- GWAS – Genome Wide Association Study (Estudo de associação genômica ampla)

SUMÁRIO

1.	INTRODUÇÃO	10
2.	OBJETIVOS.....	12
2.1	Objetivo Geral	13
2.2	Objetivos Específicos	13
3.	REVISÃO DE LITERATURA	14
3.1	Software de Simulação	14
3.2	Populações Multirraciais	14
3.3	Desequilíbrio de Ligação (DL)	15
4.	MATERIAIS E MÉTODOS	17
4.1	Simulação da população de bovinos.....	17
4.2	Estrutura da População Histórica	18
4.3	Estrutura das Populações Recentes	19
4.3	Genoma	21
4.4	População de Treinamento e Validação.....	23
5.	RESULTADOS E DISCUSSÕES	26
5.1	Desequilíbrio de Ligação	26
5.2	GWAS	29
5.3	Acurácia	31
6.	CONSIDERAÇÕES FINAIS.....	33
7.	REFERÊNCIAS BIBLIOGRÁFICAS.....	34

1. INTRODUÇÃO

A utilização de sistemas de acasalamento entre animais de diferentes raças, zootecnicamente denominados cruzamentos, constitui uma prática amplamente difundida e utilizada na criação de animais de produção mundialmente. O objetivo desse manejo reprodutivo é fazer uso da heterose e complementariedade, fenômenos genéticos inerentes aos cruzamentos e que trazem benefícios à produção. A heterose, também denominada vigor híbrido, caracteriza-se pela superioridade de média dos filhos híbridos em relação à média dos pais puros. A complementariedade trata-se de tentativa de unir, em um animal, características desejáveis que se encontravam demonstradas separadamente em uma e outra raça, através do acasalamento entre elas.

O rebanho brasileiro de bovinos de corte é composto basicamente por animais de origem asiática, segundo a Associação Brasileira das Indústrias Exportadoras de Carne (2014), oitenta por cento do rebanho comercial gado de corte brasileiro é gado zebu ou mestiço de zebu. Entretanto, visando à melhoria da qualidade da carne e carcaças, nas últimas duas décadas o Brasil está experimentando um processo de introdução de raças europeias, seja para formação de rebanhos puros ou para o cruzamento com o gado zebuino. Ademais, recentemente tem havido um esforço na formação de grupamentos genéticos que aliem as características de adaptabilidade e rusticidade já adquiridas pelo gado zebuino brasileiro com as características e produção e qualidade de carne, inatas ao gado europeu. Tais grupamentos de animais são chamados bovinos compostos, sendo o Montana Tropical, Brangus e Canchim exemplos de bovinos compostos desenvolvidos no Brasil.

A seleção genômica é a mais moderna tecnologia no que tange a utilização de marcadores genéticos para a seleção e melhoramento genético de animais domésticos. Em síntese a metodologia consiste em uma seleção assistida por marcadores em uma escala ampla do Genoma e foi proposta por Meuwissen et al. (2001). O desequilíbrio de ligação (DL) entre os marcadores e loci de características quantitativas (QTL) e a composição da população referência são pontos chave para a confiabilidade da seleção genômica. À medida que os indivíduos se distanciam geneticamente, o DL entre marcadores e QTL diminui, o que dificulta a aplicação da seleção genômica em populações de animais multirraciais e explica o fato de até hoje a maior parte das pesquisas se concentrarem na estimação de valores genômicos

apenas para animais de raças puras. O grande trunfo da seleção genômica é possibilitar a predição dos valores genômicos de candidatos à seleção jovens, ao contrário da seleção genética, na qual o candidato à seleção deve ter progênie formada para estimação do seu valor genético. Selecionar animais jovens comprovadamente superiores leva à diminuição do intervalo de gerações e consequentemente aumenta o ganho genético anual dos programas de melhoramento.

A ampliação do uso da seleção genômica para além das populações puras, ou seja, a aplicação dessa metodologia em populações multirraciais de bovinos constitui uma possibilidade real de grandes avanços no melhoramento genético do rebanho brasileiro de bovinos de corte.

2. HIPÓTESE

Compor a população de treinamento com animais de raças puras parentais permite obter boa acurácia na estimação do valor genômico de animais multirraciais para seleção genômica em bovinos de corte.

3. OBJETIVOS

3.1 Objetivo Geral

O trabalho tem por objetivo determinar a aplicabilidade da seleção genômica em populações multirraciais de animais, para tanto, a acurácia do valor genômico estimado será avaliada em diversos cenários simulados, oriundos de diferentes níveis dos seguintes fatores de variação: níveis de herdabilidade, quantidade de QTLs, composição da população de treinamento e composição da população de validação.

3.2 Objetivos Específicos

- Avaliar a melhor composição da população de treinamento para detectar SNPs a serem utilizados em animais multirraciais;
- Avaliar a correlação entre a herdabilidade da característica e a persistência dos marcadores através das gerações;
- Comparar acurácia da seleção genômica em situações que animais de grupamentos genéticos diferentes compõem a população de treinamento e a população de validação.

4. REVISÃO DE LITERATURA

4.1 Software de Simulação

Embora, nas últimas décadas, o desenvolvimento e oferta comercial de “SNP chips” de diversas densidades, por diferentes empresas, tenham diminuído substancialmente os custos relacionados à genotipagem, os estudos de seleção genômica ampla ainda são economicamente dispendiosos aos pesquisadores. Segundo Carvajal-Rodríguez (2008), simulações são ferramentas essenciais para avaliar propriedades amostrais de qualquer estatística usada em estudos de associação genômica ampla e para comparar a performance de métodos aplicados em amplas escalas de genoma. Assim sendo, foram desenvolvidos softwares de simulação voltados a estudos de populações, tais como GENESYS (EUCLYDES, 1996), FREGENE (HOGGART, 2007), GENOME (LIANG et al., 2007), entre outros. Uma vez que a grande maioria desses softwares foi desenvolvida para estudos em humanos, é necessário lembrar que simulação e estudo de dados genômicos em animais de produção se difere em vários aspectos das análises humanas. Por exemplo, Andersson (2001) cita que para detecção de QTL em animais de produção, estratégias comumente utilizadas são pedigrees com muitas gerações, grandes famílias, nas quais foi feito uso de inseminação artificial, desenhar um programa de cruzamentos, como F2 e retro-cruzamentos. Segundo Sargolzaei e Schenkel (2009), uma diferença ainda mais importante é que as populações humanas tem experimentado uma expansão no tamanho efetivo da população, enquanto o tamanho efetivo da população de animais de produção tem diminuído. Desse modo, o software QMSim (SARGOLZAEI; SCHENKEL, 2009) foi desenvolvido para simular uma extensa gama de arquiteturas genéticas e estruturas populacionais de animais de produção.

4.2 Populações Multirraciais

De acordo com Dickerson (1969), uma das maneiras para se utilizar da diversidade genética existente entre as raças para a otimização da produção animal, é fazer uso dos cruzamentos entre raças. Segundo o mesmo autor algumas razões para utilização de cruzamentos são aproveitar a heterose, complementariedade, dar base para formação de novas raças e flexibilizar os sistemas de produção.

Segundo Elzo e Borjas (2004), populações multirraciais são caracterizadas como aquelas populações compostas pela reprodução cruzada de animais de raças puras e seus cruzamentos.

Nos últimos anos é notável o crescimento da utilização de cruzamentos na pecuária de corte, isso se deve à popularização das técnicas de inseminação artificial, bem como ao interesse dos produtores em aumentar a eficiência e produtividade.

4.3 Desequilíbrio de Ligação (DL)

Desequilíbrio de ligação é a associação não aleatória de alelos entre dois loci. Trata-se de um conceito estatístico que mede o grau de associação entre alelos de diferentes loci. Quando a frequência de um determinado haplótipo (dois alelos de diferentes loci no mesmo cromossomo) é maior do que o esperado, os loci estão em desequilíbrio de ligação (ARDLIE et al., 2002).

O desequilíbrio de ligação é resultado predominantemente do tamanho efetivo da população (BRITO, 2011), comparando humanos a animais domésticos, cujos tamanhos efetivos da população são aproximadamente 10.000 e menor do que 100 (KRUGLYAK et al., 1999; RIQUET et al., 1999), níveis significantes de DL se estendem por distâncias menores que 0,005 cM em humanos (DUNNING et al., 2000; REICH et al., 2001), ao passo que em bovinos esses níveis podem se estender a 5-10 cM (HAYES et al., 2003; de ROOS et al., 2008; SARGOLZAEI et al., 2008).

O desequilíbrio de ligação ocorre quando dois genes estão localizados à uma distância mínima, de modo que, durante a meiose, a recombinação entre eles seja rara, conseqüentemente, esses dois genes (bem como os genes localizados entre eles) serão transmitidos juntos através de algumas gerações.

A mensuração mais usual de DL é dada por r^2 (HILL e ROBERTSON, 1968), por ser a medida menos sensível à frequência alélica e ao tamanho da amostra na estimação do DL entre marcadores bialélicos, tais como os SNP's (BOHMANOVA et al., 2010).

O r^2 pode ser calculado pela equação:

$$r^2 = \frac{freq(A1_B1) * freq(A2_B2) * freq(A1_B2) * freq(A2_B1)}{freq(A1) * freq(A2) * freq(B1) * freq(B2)}$$

Onde a $\text{freq}(A1_B1)$ é a frequência do haplótipo $A1_B1$ na população e $\text{freq}(A1)$ é a frequência do alelo $A1$.

Os valores de r^2 para um par de loci variam de 0 a 1, não havendo nenhum DL entre eles e quando estão em completo desequilíbrio, respectivamente.

Em bovinos, segundo Brito (2011), os estudos para descrição da extensão e padrão de DL se desenvolveram a medida que se disponibilizava painéis de marcadores mais densos.

Seleção Genômica

O termo seleção genômica foi proposto por Meuwissen et al. (2001), nessa metodologia, os genes que controlam as características de interesse (QTL), não são conhecidos, em vez disso, marcadores genéticos são relacionados à expressão da característica. A soma dos efeitos dos marcadores pode ser um preditor acurado do mérito genético dos candidatos à seleção. Painéis suficientemente densos, cobrindo todo o genoma, permitem que sejam identificados marcadores que capturem a variabilidade associada aos QTL's que influenciam características de importância econômica (NEVES, 2013).

Para tanto, faz-se necessário que cada QTL esteja em desequilíbrio de ligação com pelo menos um marcador ou haplótipo de marcadores (HAYES, 2007), logo, a densidade de marcadores é essencial para a acurácia da predição do valor genômico através de gerações e/ou populações (CALUS, 2010; HAYES et al., 2009). Quanto menor a extensão do desequilíbrio de ligação em uma população, mais denso deve ser o painel de marcadores para que a seleção genômica funcione (de ROSS et al., 2008).

5. MATERIAIS E MÉTODOS

5.1 Simulação da população de bovinos

Foi utilizado o software QMSIM (SARGOLZAEI; SCHENKEL, 2009) para a simulação das populações de bovinos de corte. O programa, desenvolvido por pesquisadores ligados à Universidade de Guelph, é baseado na linguagem de programação C++ e foi executado em computador lotado à Sala de Processamento do Núcleo de Apoio à Pesquisa em Melhoramento Genético Animal, Biotecnologia e Transgenia da Faculdade de Zootecnia e Engenharia de Alimentos da Universidade de São Paulo. O servidor utiliza sistema operacional Linux, processador Intel® Xeon® E5, memória RAM de 250GB.

As características simuladas não necessariamente mimetizam alguma característica real encontradas em bovinos de corte, apenas para efeito de simulação, diferem entre si pela herdabilidade e número de QTLs que as controlam. Os dois níveis de herdabilidade (0,15 e 0,35) combinados com duas quantidades de QTLs (63 e 1000), resultam em quatro características que constituem os quatro cenários de simulação. Todas as características têm variância fenotípica igual a um.

Tabela 1 – Cenários de simulação, de acordo com a herdabilidade e nº de QTLs.

Cenários	h^2	QTLs	Var. Fenotípica
1 - LQlh2	0,15	63	1,0
2 - LQmh2	0,35	63	1,0
3 - HQlh2	0,15	1.000	1,0
4 - HQmh2	0,35	1.000	1,0

Fonte: Autoria própria

Para simulação da população multirracial, objetivo de estudo deste trabalho, primeiramente gerou-se uma população histórica que remontava ao período anterior à diferenciação dos bovinos. A partir dessa população histórica divergiram duas populações, dois grupos de animais oriundos das últimas gerações da população históricas foram separados e continuaram se reproduzindo isoladamente, simulando o isolamento geográfico e domesticação, tal como se especula que tenha ocorrido com a espécie *Bos primigenius*, resultando no surgimento das subespécies *Bos taurus taurus* e *Bos taurus indicus* (BAILEY et

al, 1996; FAO, 2006; LOFTUS et al, 1994). Finalmente a população multirracial foi formada a partir do cruzamento dessas duas populações divergentes.

5.2 Estrutura da População Histórica

Foi simulada uma população histórica partindo de 750 machos e 750 fêmeas na geração inicial, permanecendo durante mil gerações com tamanho constante de 1500 animais, que finalmente foram reduzidos a 400 animais, destes 250 machos, na milésima septingentésima trigésima (1730ª) geração. A diminuição de 1500 animais para 400 animais em 730 gerações é importante para gerar deriva genética e iniciar o histórico de desequilíbrio de ligação, pelo fenômeno denominado gargalo genético ou *bottleneck* na língua inglesa. Essa etapa teve por objetivo simular uma população de *Bos primigenius* que nos passos seguintes da simulação dará origens às subespécies modernas de bovinos.

5.3 Estrutura das Populações Recentes

No próximo passo da simulação, 400 animais das últimas gerações da população histórica foram separados em dois grupos, formados por 100 touros e 100 vacas. Uma vez isolados, cada grupo passou por 250 gerações de união aleatória dos gametas, cada vaca gerou 5 progênies, com o intuito de expandir a população. Essa etapa mimetiza o momento em que duas subpopulações de *Bos primigenius* se isolam geograficamente e evoluem para as subespécies *Bos taurus* e *Bos taurus indicus* e aumentam naturalmente o tamanho das respectivas populações.

Segundo Henson (1992), o processo de domesticação leva a redução do tamanho efetivo da população, portanto, na tentativa de simular o processo de domesticação, ambas as subpopulações (doravante denominadas Raça 1 e Raça 2) passaram por uma redução drástica do número de animais e tamanho efetivo da população. Foram amostrados aleatoriamente 50 machos e 3000 fêmeas para a Raça 1 e 100 machos e 3000 fêmeas para a Raça 2, nos dois casos, todos os animais eram provenientes da ducentésima quinquagésima geração (250ª). Nas 20 gerações relacionadas ao período de domesticação, cada raça teve tamanho efetivo da população diferente, a primeira com 197 e a segunda com 387, com o objetivo de gerar duas populações que divergissem quanto à extensão do desequilíbrio de ligação, caracterizando duas raças diferentes de bovinos modernos, a exemplo do que foi encontrado por Lu et al (2012).

Posteriormente, foi realizado o cruzamento entre as duas raças. Cada raça contribuiu com 250 machos e 500 fêmeas, o que resultou em 1000 animais na primeira geração de cruzamento, que continuaram acasalando aleatoriamente entre si por 10 gerações para formar a população final de animais multirraciais.

Figura 1 – Estrutura Populacional.

Fonte: Autoria própria

5.4 Genoma

O genoma dos bovinos domésticos é composto por 30 pares de cromossomos, sendo 29 pares autossômicos e um par de cromossomos sexuais. O genoma bovino foi simulado com todos os cromossomos autossômicos e buscou ser o mais próximo possível da realidade, o comprimento foi idêntico ao do genoma bovino real, totalizando 2333 cM, de acordo com a montagem Btau_3.1 (Snelling et al, 2007). Foram utilizados 57.024 SNPs e, de acordo com o cenário, 70 ou 1.000 QTLs distribuídos de maneira uniforme através do genoma. Os efeitos dos QTLs foram simulados em distribuição gamma com parâmetro de forma igual a 0,40 (HAYES; GODDARD, 2001). A taxa de mutação recorrente para SNPs e QTLs simulada foi de 1×10^{-4} . Os parâmetros detalhados utilizados na simulação constam nas Tabelas 2 e 3.

Em estudos de simulação torna-se possível obter o Valor Genético Verdadeiro (VGV) dos animais para a característica em estudo. A somatória dos efeitos dos QTLs de cada animal para cada característica foi designada como o Valor Genético Verdadeiro, desse modo:

$$VGV_i = \sum_{j=1}^n (x_{ijk} g_{jk})$$

onde x_{ijk} é o número de cópias do alelo k que o indivíduo i tem na posição j e g_{jk} é o efeito do QTL do alelo k na posição j (Nascimento, 2014).

Tabela 2 – Parâmetros utilizados para simulação das populações de Bovinos de Corte.

Estrutura da população	Parâmetros	
População histórica (HP)		
Nº de animais (Nº de gerações constantes)	1500(1000)	
Nº de animais (Nº de gerações - gargalo genético)	400(730)	
Isolamento e Expansão das populações (EXP)	Subspécie 1	Subspécie 2
Nº de machos fundadores (HP)	100	100
Nº de fêmeas fundadoras (HP)	100	100
Nº de gerações	250	250
Nº de progênie por mãe	5	5
Domesticação e Populações recentes (RECs)	Raça 1	Raça2
Nº de machos fundadores (EXP)	50	100
Nº de fêmeas fundadoras (EXP)	3000	3000
Nº de gerações	20	20
Nº de filhos por mãe	1	1
Porcentagem de machos por geração de filhos	50%	50%
Esquemas de acasalamento	Aleatório	Aleatório
Seleção	Valor Genético Estimado	Aleatório
Método de estimação dos valores genéticos	BLUP	-
População Multirracial (PM)		
Nº de machos fundadores (Raça1, Raça 2)	500, 250	
Nº de fêmeas fundadoras (Raça 1, Raça 2)	500, 500	
Nº de gerações	3	
Nº de filhos por mãe	5	
Porcentagem de machos por geração de filhos	50%	
Esquemas de acasalamento	Aleatório	
Seleção	Aleatório	
Método de estimação dos valores genéticos	-	
Herdabilidade das características	0,15/0,35	
Variância fenotípica	1/1	
Genoma		
Nº de cromossomos	29	
Tamanho total	2333 cM	
Nº total de marcadores	57.024	
Nº total de QTL	1.000/63	
Distribuição dos QTL	Aleatória	
MAF dos marcadores		
Efeito alélico aditivo dos QTL	Distribuição Gama (0,4)	
Taxa de marcador genético perdido	0,01	
Taxa de erro de genotipagem	0,005	
Taxa de mutação recorrente do marcador	0,0001	
Taxa de mutação recorrente do QTL	0,0001	

Fonte: Autoria própria

Tabela 3 – Dados utilizados para simulação do genoma do boi doméstico, nº de marcadores SNPs e loci de efeito quantitativo por cromossoma.

Cromossomo	Comprimento (cM)	Nº de SNPs	Nº de QTLs	
1	146	3647	62	/4
2	126	3022	56	/4
3	116	2777	52	/3
4	111	2754	48	/3
5	119	2447	53	/3
6	112	2713	48	/3
7	101	2540	43	/3
8	104	2592	44	/3
9	95	2208	40	/3
10	96	2387	41	/3
11	102	2500	43	/2
12	78	1893	33	/2
13	83	2022	35	/2
14	82	1935	35	/2
15	75	1916	32	/2
16	73	1864	31	/2
17	70	1741	30	/2
18	63	1524	27	/2
19	63	1570	27	/2
20	68	1682	29	/2
21	63	1607	27	/2
22	60	1432	25	/2
23	49	1180	21	/1
24	60	1412	25	/1
25	42	1091	18	/1
26	48	1241	21	/1
27	43	1066	18	/1
28	40	1076	17	/1
29	45	1185	19	/1
Total	2333	57024	1000	/63

Fonte: Autoria própria

5.5 População de Treinamento e Validação

Para realizar a estimativa do valor genômico, cada grupamento genético (Raça 1, Raça 2 e Multirracial) entrou pelo menos uma vez como população de treinamento e como população de validação. Inicialmente as informações genotípicas e fenotípicas das populações de treinamento foram

utilizadas para estimar os efeitos dos SNPs e posteriormente, utilizando-se apenas dos genótipos das populações de validação, foi obtido o Valor Genômico Estimado. Da combinação das diferentes situações propostas, resultaram 33 conjuntos para estimativa dos efeitos dos SNPs, provenientes de 11 populações de treinamento diferentes e 3 populações de validação. Foram amostrados 4.000 animais de cada grupamento genético, dessa amostra, os 2.500 animais mais velhos foram designados para composição das populações de treinamento e os 1.500 mais jovens para a composição das populações de validação. Para identificar os melhores métodos de composição da população de treinamento para se estimar o valor genômico de uma população multirracial, foram pensadas as seguintes composições: populações de treinamento compostas somente por animais de cada grupamento genético, isolados; populações de treinamento compostas por animais provenientes de diferentes grupamentos genéticos, combinados dois a dois; populações de treinamento compostas por animais de todos os grupamentos genéticos envolvidos. Quando utilizou-se dois ou três grupamentos genéticos, houveram duas situações acerca do nº de animais: manter o mesmo nº de animal para todas as populações de treinamento, ou seja 2.500 animais, desse modo, quando dois grupamentos genéticos formavam o grupo de treino, cada um contribuía com metade dos indivíduos e quando todos os grupamentos genéticos foram utilizados, cada um dos três contribuiu com um terço; a outra situação foi a de se utilizar de todos os animais previamente utilizados nas populações de treinamento de cada grupamento genético, desse modo, quando a população de treinamento era composta por animais de dois grupos genéticos, esta totalizava 5.000 animais e quando três grupamentos genéticos foram utilizados totalizou 7.500 animais.

Além da população multirracial como população de validação, com o intuito de comparar valores de acurácia e testar uma validação cruzada, criou-se mais duas populações de validação, uma composta somente por animais da Raça 1 e outra composta somente por animais da Raça 2. Desse modo, com 11 populações de treinamento, versus 3 populações de validação, foram gerados 33 de conjuntos de populações de treinamento e validação, para cada uma das características simuladas.

Figura 2 – Estrutura e validação dos cenários de populações recentes.

Fonte: Autoria própria

6. RESULTADOS E DISCUSSÕES

6.1 Desequilíbrio de Ligação

Em todas as quatro características simuladas e em cada população, o desequilíbrio de ligação diminuiu, conforme o espaço entre os marcadores aumentou. A raça 1 apresentou os maiores valores de LD, em todos os casos, conforme esperado, por ter sido simulada com o tamanho efetivo da população menor, em relação à raça 2, que por sua vez, apresentou LD médio inferior. A população multirracial, simulada para obter valores menores de LD em relação as duas raças parentais, conforme os estudos de Lu et al (2012), no entanto, em alguns casos, apresentou LD médio maior que a raça 2. O que pode ter sido causado pela diferença no número de gerações simuladas. Ao passo que as raças 1 e 2 tiveram 20 gerações simuladas, para a população multirracial foram simuladas apenas três gerações.

Tabela 4 – Desequilíbrio de ligação médio (r^2) e desvios-padrão, na última geração de cada grupamento racial, no cenário de 63 QTLs e $h^2 = 0,15$.

	Raça 1	Raça 2	Multirracial
Distância (cM)	r^2 médio (desvio padrão)		
[0,.05)	0.4508(0.3665)	0.2834(0.2929)	0.2865(0.2917)
[.05,.1)	0.4123(0.3550)	0.2468(0.2650)	0.2523(0.2684)
[.1,.2)	0.3747(0.3416)	0.2103(0.2393)	0.2106(0.2382)
[.2,.3)	0.3380(0.3206)	0.1737(0.2053)	0.1757(0.2049)
[.3,.4)	0.3079(0.3050)	0.1468(0.1804)	0.1505(0.1815)
[.4,.5)	0.2902(0.2929)	0.1319(0.1634)	0.1338(0.1637)
[.5,.6)	0.2712(0.2806)	0.1182(0.1499)	0.1191(0.1496)
[.6,.7)	0.2681(0.2771)	0.1064(0.1368)	0.1087(0.1377)
[.7,.8)	0.2555(0.2695)	0.1000(0.1288)	0.1013(0.1292)
[.8,.9)	0.2503(0.2669)	0.0911(0.1176)	0.0939(0.1225)
[.9,1)	0.2383(0.2578)	0.0850(0.1107)	0.0878(0.1148)
[1,2)	0.2157(0.2379)	0.0643(0.0867)	0.0662(0.0888)
[2,3)	0.1838(0.2105)	0.0442(0.0606)	0.0458(0.0630)
[3,4)	0.1684(0.1956)	0.0347(0.0484)	0.0360(0.0503)
[4,5)	0.1553(0.1827)	0.0283(0.0396)	0.0297(0.0414)

Fonte: Autoria própria

Tabela 5 – Desequilíbrio de ligação médio (r^2) e desvios-padrão, na última geração de cada grupamento racial, no cenário de 63 QTLs e $h^2 = 0,35$.

	Raça 1	Raça 2	Multirracial
Distância (cM)	r^2 médio (desvio padrão)		
[0,.05)	0.4107(0.3547)	0.2762(0.2895)	0.2810(0.2912)
[.05,.1)	0.3742(0.3423)	0.2428(0.2650)	0.2469(0.2673)
[.1,.2)	0.3364(0.3221)	0.2086(0.2388)	0.2079(0.2355)
[.2,.3)	0.2976(0.2984)	0.1719(0.2033)	0.1697(0.2001)
[.3,.4)	0.2711(0.2825)	0.1478(0.1801)	0.1434(0.1733)
[.4,.5)	0.2536(0.2692)	0.1284(0.1601)	0.1265(0.1580)
[.5,.6)	0.2399(0.2599)	0.1172(0.1468)	0.1134(0.1454)
[.6,.7)	0.2265(0.2485)	0.1057(0.1359)	0.1028(0.1318)
[.7,.8)	0.2201(0.2440)	0.0995(0.1292)	0.0962(0.1238)
[.8,.9)	0.2145(0.2377)	0.0904(0.1189)	0.0897(0.1175)
[.9,1)	0.2049(0.2306)	0.0853(0.1115)	0.0835(0.1086)
[1,2)	0.1804(0.2099)	0.0651(0.0883)	0.0632(0.0856)
[2,3)	0.1534(0.1828)	0.0450(0.0622)	0.0432(0.0596)
[3,4)	0.1379(0.1661)	0.0343(0.0478)	0.0329(0.0462)
[4,5)	0.1261(0.1542)	0.0286(0.0395)	0.0271(0.0381)

Fonte: Autoria própria

Tabela 6 – Desequilíbrio de ligação médio (r^2) e desvios-padrão, na última geração de cada grupamento racial, no cenário de 1000 QTLs e $h^2 = 0,15$.

	Raça 1	Raça 2	Multirracial
Distância (cM)	r^2 médio (desvio padrão)		
[0,.05)	0.4817(0.3741)	0.2837(0.2933)	0.2811(0.2921)
[.05,.1)	0.4457(0.3634)	0.2545(0.2727)	0.2512(0.2672)
[.1,.2)	0.4091(0.3513)	0.2160(0.2401)	0.2143(0.2392)
[.2,.3)	0.3743(0.3354)	0.1763(0.2070)	0.1785(0.2069)
[.3,.4)	0.3431(0.3206)	0.1496(0.1812)	0.1524(0.1815)
[.4,.5)	0.3257(0.3117)	0.1336(0.1640)	0.1356(0.1653)
[.5,.6)	0.3101(0.3028)	0.1196(0.1491)	0.1211(0.1498)
[.6,.7)	0.3004(0.2966)	0.1103(0.1391)	0.1112(0.1397)
[.7,.8)	0.2883(0.2884)	0.1005(0.1293)	0.1025(0.1292)
[.8,.9)	0.2775(0.2818)	0.0929(0.1201)	0.0943(0.1213)
[.9,1)	0.2711(0.2781)	0.0875(0.1137)	0.0890(0.1138)
[1,2)	0.2429(0.2564)	0.0652(0.0883)	0.0667(0.0899)
[2,3)	0.2107(0.2307)	0.0449(0.0623)	0.0461(0.0638)
[3,4)	0.1890(0.2122)	0.0348(0.0488)	0.0359(0.0498)
[4,5)	0.1743(0.1987)	0.0287(0.0402)	0.0296(0.0415)

Fonte: Autoria própria

Tabela 7 – Desequilíbrio de ligação médio (r^2) e desvios-padrão, na última geração de cada grupamento racial, no cenário de 1000 QTLs e $h^2 = 0,35$.

	Raça 1	Raça 2	Multirracial
Distância (cM)	r^2 médio (desvio padrão)		
[0,.05)	0.4034(0.3539)	0.2882(0.2951)	0.2828(0.2915)
[.05,.1)	0.3737(0.3395)	0.2518(0.2691)	0.2519(0.2659)
[.1,.2)	0.3355(0.3213)	0.2093(0.2360)	0.2125(0.2365)
[.2,.3)	0.2957(0.2969)	0.1730(0.2031)	0.1723(0.2017)
[.3,.4)	0.2722(0.2802)	0.1482(0.1772)	0.1487(0.1780)
[.4,.5)	0.2547(0.2685)	0.1316(0.1615)	0.1303(0.1608)
[.5,.6)	0.2428(0.2606)	0.1179(0.1478)	0.1178(0.1475)
[.6,.7)	0.2246(0.2463)	0.1094(0.1374)	0.1073(0.1363)
[.7,.8)	0.2183(0.2407)	0.0987(0.1274)	0.0975(0.1255)
[.8,.9)	0.2081(0.2321)	0.0913(0.1183)	0.0898(0.1152)
[.9,1)	0.2037(0.2291)	0.0864(0.1121)	0.0856(0.1109)
[1,2)	0.1788(0.2060)	0.0653(0.0879)	0.0643(0.0866)
[2,3)	0.1515(0.1798)	0.0453(0.0622)	0.0438(0.0601)
[3,4)	0.1348(0.1633)	0.0348(0.0485)	0.0332(0.0464)
[4,5)	0.1242(0.1514)	0.0291(0.0411)	0.0272(0.0387)

Fonte: Autoria própria

6.2 GWAS

Em todos os cenários foi utilizada a mesma densidade de marcadores, porém, em metade dos cenários utilizou-se propositalmente alta e baixa quantidade de QTLs (63 e 1000, respectivamente). Desse modo, como era de se esperar, um número muito maior de marcadores estava em desequilíbrio de ligação com os QTLs nos cenários em que esses estavam presentes em maior quantidade, o que nos *Manhattan Plots* pode ser visto observando a grande quantidade de regiões cromossômicas com grande participação na variância da característica. Por outro lado, nos cenários em que se utilizou baixa quantidade de QTLs, menos regiões cromossômicas tem tanta influência na variância da características, o que faz com que regiões cromossômicas importantes apresentem pico mais evidente.

Figura 3 - Comparação dos *Manhattan Plots* dos diferentes cenários

Fonte: Autoria própria

6.3 Acurácia

Os valores de acurácia obtidos em cada um dos quatro cenários estão demonstrados nas tabelas 3, 4, 5 e 6. Cada uma delas apresenta os 11 cenários de populações de treinamento formadas, combinadas com as três populações de validação.

Quando a mesma raça pura foi utilizada como população de treinamento e validação, obteve-se boa acurácia, em todos os cenários, variando de 0,45 (raça 2, 1000 QTLs e 0,15 de h^2) a 0,68 (raça1, 63 QTLs e 0,30 de h^2).

Usando animais da população multirracial para formação de ambas as populações, treinamento e validação, também foram obtidos valores de boa acurácia, variando de 0,65 (1000 QTLs, 0,15 de h^2) a 0,80 (63 QTLs e 0,15 de h^2).

Tabela 8 – Acurácia da predição do Valor Genômico Estimado utilizando 1.000 QTLs e $h^2 = 0,30$.

1.000 QTLs, $h^2 = 0,30$ Pop. De Treinamento	Pop. de Validação		
	Raça 1	Raça 2	Multirracial
Raça1	0.6368792	0.2876386	0.5980408
Raça2	0.3549584	0.6368792	0.4216392
Multirracial	0.3964652	0.2593261	0.7142194
Raça1 + Raça2 (n=2500)	0.6086603	0.4451556	0.7753878
Raça1 + Raça2 (n=5000)	0.698937	0.5876915	0.8054714
Raça1 + Multirracial (n=2500)	0.4087272	0.1367461	0.5993641
Raça1 + Multirracial (n=5000)	0.5499597	0.3327992	0.2624072
Raça2 + Multirracial (n=2500)	0.4586738	0.4706239	0.7983068
Raça2 + Multirracial (n=5000)	0.5430871	0.6268421	0.8494629
Raça1 + Raça2 + Multirracial (n=2500)	0.4185263	0.4000995	0.6851689
Raça1 + Raça2 + Multirracial (n=5000)	0.5918543	0.5608263	0.7318141

Fonte: Autoria própria

Tabela 9 – Acurácia da predição do Valor Genômico Estimado utilizando 1.000 QTLs e $h^2 = 0,15$.

1.000 QTLs, $h^2 = 0,15$	Pop. de Validação		
Pop. De Treinamento	Raça 1	Raça 2	Multirracial
Raça1	0.577415	0.1623812	0.477355
Raça2	0.1135576	0.4501537	0.3856592
Multirracial	0.381634	0.3025857	0.6541515
Raça1 + Raça2 (n=2500)	0.4711122	0.3736845	0.6828902
Raça1 + Raça2 (n=5000)	0.6047857	0.4873567	0.7142184
Raça1 + Multirracial (n=2500)	0.3102862	0.159486	0.4272053
Raça1 + Multirracial (n=5000)	0.5183503	0.2691102	0.05799612
Raça2 + Multirracial (n=2500)	0.3822822	0.4084424	0.70333
Raça2 + Multirracial (n=5000)	0.444206	0.5624557	0.7780636
Raça1 + Raça2 + Multirracial (n=2500)	0.4490039	0.4094913	0.6625324
Raça1 + Raça2 + Multirracial (n=5000)	0.5640012	0.5404541	0.7183813

Fonte: Autoria própria

Tabela 10 – Acurácia da predição do Valor Genômico Estimado utilizando 63 QTLs e $h^2 = 0,30$.

63 QTLs, $h^2 = 0,30$	Pop. de Validação		
Pop. De Treinamento	Raça 1	Raça 2	Multirracial
Raça1	0.6861241	0.1768667	0.5588474
Raça2	0.2186276	0.5288755	0.3078796
Multirracial	0.5403914	0.2443698	0.6862314
Raça1 + Raça2 (n=2500)	0.5795235	0.4118927	0.6893869
Raça1 + Raça2 (n=5000)	0.6899473	0.5449797	0.7167666
Raça1 + Multirracial (n=2500)	0.6115305	0.2407342	0.6734686
Raça1 + Multirracial (n=5000)	0.7224122	0.2955408	0.7224639
Raça2 + Multirracial (n=2500)	0.4824702	0.4445019	0.6858906
Raça2 + Multirracial (n=5000)	0.5541383	0.5677548	0.7254326
Raça1 + Raça2 + Multirracial (n=2500)	0.5996665	0.3867474	0.7031371
Raça1 + Raça2 + Multirracial (n=5000)	0.7257098	0.5768587	0.7520037

Fonte: Autoria própria

Tabela 11 – Acurácia da predição do Valor Genômico Estimado utilizando 63 QTLs e $h^2 = 0,15$.

63 QTLs, $h^2 = 0,15$		Pop. de Validação		
Pop. De Treinamento	Raça 1	Raça 2	Multirracial	
Raça1	0.4827135	0.1465097	0.4890996	
Raça2	0.0460818	0.5338164	0.3513411	
Multirracial	0.4388386	0.4264237	0.8069339	
Raça1 + Raça2 (n=2500)	0.4977754	0.4361726	0.7669976	
Raça1 + Raça2 (n=5000)	0.5764612	0.4243295	0.8277802	
Raça1 + Multirracial (n=2500)	0.5314726	0.336591	0.7377936	
Raça1 + Multirracial (n=5000)	0.5764612	0.4243295	0.8277802	
Raça2 + Multirracial (n=2500)	0.3754404	0.4624386	0.7392575	
Raça2 + Multirracial (n=5000)	0.444323	0.5880381	0.8272432	
Raça1 + Raça2 + Multirracial (n=2500)	0.4368685	0.4272738	0.7538077	
Raça1 + Raça2 + Multirracial (n=5000)	0.5936343	0.5981002	0.8423029	

Fonte: Autoria própria

7. CONSIDERAÇÕES FINAIS

Softwares de simulação são importantes ferramentas no contexto de estudo e pesquisa da genômica, porém, sua ficha de parâmetros é ampla e o usuário deve estar atento às implicações de cada um dele nos dados gerados a partir da simulação.

Novas simulações devem ser feitas, para que se possa cumprir os objetivos deste trabalho, discutir corretamente os resultados e realizar as devidas conclusões.

8. REFERÊNCIAS BIBLIOGRÁFICAS

ANDERSSON, L. **Genetic dissection of phenotypic diversity in farm animals**. Nature Genetics Reviews 2, 130-138, 2001.

ARDLIE, K.G.; KRUGLYAK, L.; SEIELSTAD, M. **Patterns of linkage disequilibrium in the human genome**. Nature Reviews: Genetics, London, v.3, p. 299-309, 2002.

BAILEY, J.F. et al. **Ancient DNA suggests a recent expansion of European cattle from a diverse wild progenitor species**. Proc R Soc Lond B Biol Sci; 263:1467, 1996.

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

Bovine HapmapConsortium. **Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds**. Science 324:528-532 (2009).

CALUS, M.P.L. **Genomic breeding value prediction: methods and procedures**. Animal, Cambridge, v. 4, n. 2, p 157-164, 2010.

CARVAJAL-RODRÍGUEZ, A. **Simulation of Genomes: A Review**. Current Genomics, Vol. 9, No. 3, 2008.

DE ROSS, A.P.W. et al. **Linkage disequilibrium and persistence of phase in Holstein Friesian, Jersey and Angus cattle**. Genetics, Bethesda, v.179, n. 3, p. 1503-1512, 2008.

DICKERSON, G. E. **Experimental approaches in utilizing breed resources**. Animal Breeding Abstracts, Wallingford, UK, 37(2):191-202, 1969.

DUNNING, A.M. et al. **The extent of linkage disequilibrium in four populations with distinct demographic histories.** Am. J. Hum. Genet., Amsterdam, v.67, p.1544-1554, 2000.

ELZO, M.A.; BORJAS, A.R. **Perspectivas da avaliação genética multirracial em bovinos no Brasil.** Ciência Animal Brasileira v.5, n.4, p. 171-185, 2004.

EUCLYDES, R. F. **Genesys: sistema de simulação genética.** Viçosa: UFV, 1996.

Food and Agriculture Organization - FAO. **The state of the world's animal genetic resources for food and agriculture - first draft.** Roma; 2006.

HAYES, B.J. et al. **Invited review: Genomic selection in dairy cattle: Progress and challenges.** Journal of Dairy Science, Champaign, v. 92, n. 3, p. 433–443, 2009.

HAYES, B.J. et al. **Novel multilocus measure of linkage disequilibrium to estimate past effective population size.** Genome Res., Woodbury, v.13, p.635-643, 2003.

HAYES, B.; GODDARD, M. E. **The distribution of the effects of genes affecting quantitative traits in livestock.** Genetics Selection Evolution, v. 33, n.3, p.209-229, 2001.

HILL, W.G.; ROBERTSON, A. **Linkage disequilibrium in finite populations.** Theor. Appl. Genet., Stuttgart, v.38, p.226-231, 1968.

HENSON, E. L. **In situ Conservation of Livestock and Poultry.** In: (2nd ed.),FAO Animal Production and Health, 1992.

HOGGART, C.J. et al. **Sequence-level population simulations over large genomic regions.** Genetics, v.177, p.1725-1731. 2007.

KRUGLYAK, L. **Prospects for whole-genome linkage disequilibrium mapping of common disease genes.** Nature Genetics, London, v.22, p.139-144, 1999.

LIMING, L. et al. **GENOME: a rapid coalescent-based whole genome simulator**. Bioinformatics 23(12):1565-7, 2007.

LOFTUS R.T. et al. **Evidence for two independent domestications of cattle**. Proc Natl Acad Sci USA, 91:2757-61, 1994.

LU D. et al. **Linkage disequilibrium in Angus, Charolais, and Crossbred beef cattle**. Frontiers in Genetics; 3:152, 2012. doi:10.3389/fgene.2012.00152.

MISZTAL I. et al. **Blupf90 and related programs (BGF90)**. Proc. 7th World Congress on Genetics Applied to Livestock Production, Montpellier, France, Commun. No. 28–07, 2002.

NASCIMENTO, G. B. **Efeito da endogamia na seleção genômica em populações simuladas de aves poedeiras**. Jaboticabal, 2014 viii, 42 p. : il. ; 28 cm.

NEVES, H.H.R. **Strategies to improve the efficiency of genomic selection in animal breeding programs**. Jaboticabal, 162 p. : il., 2013.

PURCELL, S. et al. **PLINK: a toolset for whole-genome association and population-based linkage analysis**. American Journal of Human Genetics, 81, 2007.

REICH, D.E.; LANDER, E.S. **On the allelic spectrum of human disease**. Trends Genet., Amsterdam, v.17, p.502-510, 2001.

RIQUET, J. et al. **Identity-by-descent fine- mapping of QTL in outbred populations: application to milk production in dairy cattle**. Proc. Natl. Acad. Sci., Washington, v.96, p.9252-9257, 1999.

SARGOLZAEI, M. et al. **Extent of Linkage Disequilibrium in Holstein Cattle in North America**, J. Dairy Sci., Champaign, v.91, p.2106-2117, 2008.

SARGOLZAEI, M.; SCHENKEL, F. S. **QMSim: a large-scale genome simulator for livestock**. Bioinformatics, v. 25, p. 680–681, 2009.

SNELLING, W.M. et al for the International Bovine BAC Mapping Consortium: **A physical map of the bovine genome**. Genome Biology, 8:R165, 2007.