

UNIVERSIDADE ESTADUAL PAULISTA - UNESP
CÂMPUS DE JABOTICABAL

AVALIAÇÃO DA AUTOZIGOSIDADE EM VACAS NELORE
(*Bos indicus*) ATRAVÉS DE GENÓTIPOS *SNP* DE ALTA
DENSIDADE

Ludmilla Balbo Zavarez

Médica Veterinária

2014

UNIVERSIDADE ESTADUAL PAULISTA - UNESP
CÂMPUS DE JABOTICABAL

AVALIAÇÃO DA AUTOZIGOSIDADE EM VACAS NELORE
(*Bos indicus*) ATRAVÉS DE GENÓTIPOS SNP DE ALTA
DENSIDADE

Ludmilla Balbo Zavarez

Orientador: Prof. Adj. José Fernando Garcia

Dissertação apresentada à Faculdade de Ciências Agrárias e Veterinárias - Unesp, Campus de Jaboticabal, como parte das exigências para obtenção do título de Mestre em Medicina Veterinária (Reprodução Animal).

DADOS CURRICULARES DO AUTOR

LUDMILLA BALBO ZAVAREZ - filha de Ivone Aparecida Balbo Zavarez e João Edno Zavarez, nascida em 05 de julho de 1988 em Araçatuba-SP, Brasil. Zavarez é formada pelas Faculdades Adamantinenses Integradas (FAI), como Médica Veterinária (Bacharelado em Medicina Veterinária) em 2011. Nesse mesmo ano, realizou estágio no Laboratório de Bioquímica e Biologia Molecular Animal (LBBMA) na Universidade Estadual Paulista (UNESP), Campus de Araçatuba, sob a supervisão do Prof. Adj. José Fernando Garcia e no Departamento de Reprodução Animal na Universidade Estadual Paulista (UNESP), Campus de Jaboticabal, sob a supervisão do Prof. Dr. Joaquim Mansano Garcia. Iniciou seu mestrado por meio do Programa de Pós-Graduação em Medicina Veterinária da FCAV-UNESP, Campus de Jaboticabal, em agosto de 2012, com foco na área de Reprodução Animal. Durante o período de mestrado Zavarez recebeu bolsas de estudos da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP). No ano de 2013, Zavarez cursou disciplinas de pós-graduação em dois campi da Unesp, Araçatuba e Jaboticabal. Dedicou-se também ao desenvolvimento do projeto de mestrado e participou na publicação de artigos juntamente com membros do grupo de pesquisa do LBBMA. De Fevereiro a Junho de 2014, Zavarez realizou estágio de pesquisa no exterior, no Departamento de Agricultura dos Estados Unidos (*United States Department of Agriculture - USDA*), sob supervisão do Pesquisador Dr. Tad S. Sonstegard.

“Se alguém jamais cometeu um erro é porque nunca fez nada de novo”.

Albert Einstein

*“No que diz respeito ao empenho, ao compromisso, ao esforço, à dedicação, não existe meio termo.
Ou você faz uma coisa bem feita ou não faz”.*

Ayrton Senna

“O fracasso é uma glória reservada aos fortes que se atreveram a conquistá-la”.

Autor desconhecido

DEDICO

Aos meus pais (João e Ivone) e irmãos (Larissa e Vitor) pelo incentivo, apoio e amor incondicional.

Amo vocês!

AGRADECIMENTOS

São meus sinceros agradecimentos....

Primeiramente a Deus pela conquista e realização de mais uma etapa da vida!

Ao meu orientador Prof. Adjunto José Fernando Garcia, pela orientação, os ensinamentos valiosos e exemplo de excelente pesquisador. Agradeço também pela oportunidade de desenvolvimento desta pesquisa e por ter contribuído no meu crescimento profissional, como também por toda dedicação, apoio e paciência durante o desenvolvimento dos nossos trabalhos.

Ao Prof. Dr. Joaquim Mansano Garcia pela oportunidade de ingressar nessa instituição desde os estágios extracurriculares na época da graduação, por ter me aberto as portas do seu laboratório juntamente com a Marina, meu enorme carinho!

Ao João, companheiro de todos os momentos durante o curso, por me fortalecer, estimular, apoiar, ter paciência, dedicação e amor. Te amo!

Aos meus avós, tios e primos por serem meus amigos, companheiros e acima de tudo, minha família. Obrigada por tudo!

À minha amiga de pós-graduação Roberta pela amizade, companherismo e carinho.

Ao meu caro colega de laboratório Yuri Tani Utsunomiya pelos ensinamentos, pela ajuda imensurável no desenvolvimento deste trabalho, pelo apoio e paciência.

À todos os colegas do Laboratório de Bioquímica e Biologia Molecular Animal da Unesp de Araçatuba, por todos os momentos de convivência: Prof^a Cáris, Pier, Anirene, Adam, Silvana, Fernanda, Luís, Dandara, Flávia, Samanta, Guilherme e Rafael.

Ao Dr. Tad Sonstegard pela colaboração e oportunidade de trabalhar no seu laboratório e os demais pesquisadores do Departamento de Agricultura dos Estados Unidos – USDA.

Aos membros da banca do exame de qualificação, Dr. Roberto Carneiro e Prof. Dr. Fernando Sebastian Baldi Rey pelas excelentes sugestões e contribuições na correção deste trabalho.

À Conexão DeltaGen pela disponibilização dos animais.

À todos os funcionários da Seção de Pós-graduação da Unesp de Jaboticabal e Araçatuba.

À CAPES e a FAPESP pelo apoio e incentivo à pesquisa.

Há tantos a agradecer, por tanto se dedicaram a mim, não somente por terem ensinado, mas por terem me feito aprender! A palavra mestre, nunca fará justiça aos professores dedicados, aos quais, sem nominar terão o meu eterno agradecimento.

Enfim, a todos que contribuíram de alguma maneira para minha formação, seja por meio de ensinamentos ou por momentos e palavras de amizade.

MUITO OBRIGADA!

Ludmilla Balbo Zavarez

Jaboticabal, Dezembro 2014.

APOIO FINANCEIRO

Este projeto obteve bolsa do país concedida pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), sob processo nº 2013/15869-2 e Bolsa Estágio de Pesquisa no Exterior (BEPE) nº 2013/26269-6.

SUMÁRIO

RESUMO.....	IV
TERMOS-CHAVE	IV
ABSTRACT.....	V
LISTA DE ABREVIATURAS.....	VI
LISTA DE TABELAS	IX
LISTA DE FIGURAS	X
 CAPÍTULO 1 – A AUTOZIGOZIDADE GENÔMICA E SEU POTENCIAL NO MELHORAMENTO GENÉTICO EM BOVINOS NO BRASIL	 1
1. INTRODUÇÃO	1
2. REVISÃO DE LITERATURA.....	2
3. REFERÊNCIAS	13
 CAPÍTULO 2 - ASSESSMENT OF AUTOZYGOSITY IN NELLORE COWS (<i>Bos indicus</i>) THROUGH HIGH-DENSITY <i>SNP</i> GENOTYPES	 19
ABSTRACT	20
1. INTRODUCTION.....	21
2. MATERIALS AND METHODS.....	22

2.1. Ethical statement	22
2.2. Genotyping and data filtering	22
2.3. Estimates of genomic individual autozygosity	22
2.4. Detection of common runs of homozygosity	24
3. RESULTS AND DISCUSSION.....	25
3.1. DISTRIBUTION OF <i>ROH</i> LENGTH	25
3.2. DISTRIBUTION OF GENOME-WIDE AUTOZYGOSITY	25
3.3. DISTRIBUTION OF CHROMOSOME-WISE AUTOZYGOSITY	27
3.4. IDENTIFICATION OF COMMON <i>ROH</i>	28
4. CONCLUSIONS.....	30
5. ACKNOWLEDGEMENT	31
6. REFERENCES.....	31
CAPÍTULO 3 – CONSIDERAÇÕES FINAIS	38
APÊNDICES.....	40
APÊNDICE A – FIGURE 1. Frequency distribution of percentage of autosomal genome coverage by runs of homozygosity (F_{ROH}) of different minimum lengths (>0.5, >1, >2, >4, >8 and >16 Mb) and diagonal elements of the realized genomic relationship matrix ($G_{i,i}$).....	40
APÊNDICE B – FIGURE 2. Scatterplots (lower panel) and correlations (upper panel) of percentage of autosomal genome coverage by runs of homozygosity (F_{ROH}) of different minimum lengths (>0.5, >1, >2, >4, >8 and >16 mb) and diagonal elements of the realized genomic relationship matrix ($G_{i,i}$).....	41
APÊNDICE C – FIGURE 3. Barplot of average percentage of chromosome coverage by runs of homozygosity (F_{ROH}) of minimum length of 0.5 Mb.....	41

APÊNDICE D – FIGURE 4. Manhattan plot of genome-wide locus autozygosity in Nellore cows. the dashed line represents the 50% threshold.	42
APÊNDICE E – FIGURE S1. Frequency distribution of percentage of the total genome (autosomes + x) coverage by runs of homozygosity (F_{ROH}) of different minimum lengths (>0.5, >1, >2, >4, >8 and >16 Mb) and diagonal elements of the realized genomic relationship matrix ($G_{i,i}$)	43
APÊNDICE F – FIGURE S2. Scatterplots (lower panel) and correlations (upper panel) of percentage of total genome (autosomes + x) coverage by runs of homozygosity (F_{ROH}) of different minimum lengths (>0.5, >1, >2, >4, >8 and >16 Mb) and diagonal elements of the realized genomic relationship matrix ($G_{i,i}$)	44
APÊNDICE G – TABLE 1. Detection of common runs of homozygosity according to different number of consecutive <i>SNPs</i> , percentage of animals, gap size and number of heterozygous genotypes.....	45
APÊNDICE H – TABLE S1. Parameter estimates for the nonlinear regression models by chromosome.....	46
APÊNDICE I – TABLE S2. Parameter estimates for the nonlinear regression model for 5 replicates of 1 million randomly sampled r^2 values from each autosomal chromosome.....	47
APÊNDICE J – Methods for effective population size estimation.....	47

AVALIAÇÃO DA AUTOZIGOSIDADE EM VACAS NELORE (*Bos indicus*) ATRAVÉS DE GENÓTIPOS *SNP* DE ALTA DENSIDADE

RESUMO – A pecuária bovina de corte da atualidade conta com o auxílio significativo dos programas de melhoramento genético animal para seu sucesso. O desenvolvimento de distintos métodos para realizar avaliações fenotípicas e genotípicas, tem o intuito de selecionar os rebanhos e obter animais geneticamente superiores com produção padronizada. A padronização produtiva de um rebanho deve ser resultante de acasalamentos dirigidos, que podem levar ao aumento da ocorrência de alelos homozigotos no genoma dos animais. Esse aumento da autozigosidade na população, isto é, o aumento do número de animais com alelos idênticos por descendência (*IBD*), provém do acasalamento entre indivíduos aparentados. Nesse contexto, a ferramenta mais robusta para identificar os fragmentos do genoma animal que são idênticos por descendência é a estimação do coeficiente de endogamia baseado nas corridas de homozigosidade (*ROH*), que é denominado (F_{ROH}). Eventos evolutivos, tais como: o processo de seleção natural, a deriva genética ou aleatória, e o gargalo populacional, podem ter contribuído com o aumento da autozigosidade e da ocorrência de alelos *IBD* no genoma de uma população resultante da reprodução entre indivíduos aparentados. Essa dissertação revisou o estado da arte do que concerne o estudo da endogamia determinada por análises genômicas (Capítulo 1), e aplicou tais conhecimentos na análise da distribuição dos níveis de autozigosidade baseado em corridas de homozigosidade (*ROH*) em 1.278 fêmeas da raça Nelore genotipadas para mais de 777 mil *SNPs*, e identificou regiões autozigotas possivelmente associadas à seleção natural, domesticação, fertilidade, evolução e adaptação dos bovinos (Capítulo 2). O presente estudo abre perspectivas para a realização de análises detalhadas que permitam a mensuração da influência da autozigosidade genômica causada pela pressão seletiva sobre alelos específicos em bovinos da raça Nelore.

Termos-chave: Adaptação, *Bos indicus*, Corrida de Homozigosidade, Domesticação, Endogamia, Genômica Animal, Seleção, *SNP*

ASSESSMENT OF AUTOZYGOSITY IN NELLORE COWS (*Bos indicus*) THROUGH HIGH-DENSITY SNP GENOTYPES

ABSTRACT – The current beef cattle chain has had significant help of animal breeding programs to its success. The development of different methods to perform phenotypic and genotypic evaluation has the objective of selecting flocks and producing genetically superior animals with standardized production. The productive standardization of a herd must be the result of directed matings, which can lead to increased incidence of homozygous alleles in the genome of animals. The increase of autozygosity in the population, i.e., increasing the number of animals with alleles identical by descent (*IBD*), is derived from the mating of related individuals. In this context, the most robust tool to identifies the animal's genome fragments that are identical by descent is the estimation of the inbreeding coefficient based on runs of homozygosity (*ROH*), which is called (F_{ROH}). Evolutionary events, such as the process of natural selection, random or genetic drift, and population bottleneck, may have contributed towards increasing the autozygosity and occurrence of *IBD* alleles in the genome of a population resulting crosses of related individuals. This thesis reviewed the state of the art with respect to the study of inbreeding determined by genomic analysis (Chapter 1), and applied this knowledge in the analysis of the distribution of autozygosity levels based on runs of homozygosity (*ROH*) in 1,278 Nelore females genotyped to more than 777,000 SNPs, and identified autozygote regions possibly associated with natural selection, domestication, fertility, evolution and adaptation of cattle (Chapter 2). This study open prospects for conducting detailed analysis on the influence of genomic autozygosity caused by selective pressure on specific alleles in Nelore cattle.

Keywords: Adaptation, *Bos indicus*, Runs of Homozygosity, Domestication, Inbreeding, Animal Genomics, Selection, *SNP*

LISTA DE ABREVIATURAS

ABCZ: Associação Brasileira dos Criadores de Zebu; Brazilian Association of Zebu Breeders;

BOKU: Universidade de Recursos Naturais e Ciências Biológicas; University of Natural Resources and Life Sciences;

BRCA2: Câncer de mama 2 gene; Breast cancer 2 gene;

BTA: *Bos Taurus*;

cM: Centi morgans;

DE: Depressão endogâmica; Inbreeding Depression;

DG: Duração da gestação; Gestation length;

DNA: Ácido desoxirribonucleico; Deoxyribonucleic acid;

F: Coeficiente de endogamia; Inbreeding coefficient;

F_{PED} : Coeficiente de endogamia baseado na análise de pedigree;

F_{ROH} : Coeficiente de endogamia baseado nas corridas de homozigosidade (*ROH*);

F_{ST} : Coeficiente estimado por Wright, subtotal e total da população;

Gb: Gigabase (10^9 nucleotides);

HD: Alta densidade; Hight density;

HH: Altura da garupa; Hip height;

i.e.: *Id est* (latim para “isto é” ou “em outras palavras”);

IBD: Alelos idênticos por descendência; Alleles identical by descent;

IBS: Alelos idênticos por estado; Alleles identical by state;

IEP: Intervalos entre partos; Calving intervals;

IPC: Idade a primeira cria; Age of first calving;

Kb: Kilobase (10^3 nucleotídeos);

Kg: kilogramas;

MAF: Frequência do menor alelo; Minor allele frequency;

Mb: Megabase (10^6 nucleotídeos);

mtDNA: DNA mitocondrial; Mitochondrial DNA;

MUS: Escore visual para musculatura aos 18 meses; Visual score for muscles at 18 months;

P12: Peso aos 12 meses de idade; Weight at 12 months of age;

P18: Peso aos 18 meses de idade; Weight at 18 months of age;

P24: Peso aos 24 meses de idade; Weight at 24 months of age;

P8: Peso aos 8 meses de idade; Weight at 8 months of age;

PL: Produção do leite; Milk Production;

PN: Peso ao nascimento; Birth weight;

PP14: Probabilidade de prenhes em novilhas aos 14 meses; Probability of pregnant heifers at 14 months;

PMGA: Programas de melhoramento genético animal; Animal breeding programs;

ROH: Corridas de homozigosidade; Runs of Homozygosity;

S1: Material suplementar 1; Supplementary material 1;

S2: Material suplementar 2; Supplementary material 2;

SC: Circunferência escrotal aos 18 meses; Scrotal circumference at 18 months;

SNP: Polimorfismo de sítio único; *Single Nucleotide Polymorphism*;

SPOCK1: *Sparc/ Osteonectin, Cwcv and Kazal-Like Domains Proteoglycan gene (Testican)*;

STAY: Longevidade; Longevity or Stayability;

Th2: Células T auxiliaries; T helper cells;

WG345: Ganho de peso a desmama aos 345 dias; Weaning weight gain at 345 days

WW: Peso a desmama; Weaning weight;

LISTA DE TABELAS

CAPÍTULO 1 – A AUTOZIGOZIDADE GENÔMICA E SEU POTENCIAL NO MELHORAMENTO GENÉTICO EM BOVINOS NO BRASIL

Tabela 1. Fenótipos (crescimento, reprodução e produção de leite) analisados em distintas raças bovinas no Brasil, segundo distintos autores.

CAPÍTULO 2 - ASSESSMENT OF AUTOZYGOSITY IN NELLORE COWS (*Bos indicus*) THROUGH HIGH-DENSITY *SNP* GENOTYPES

Table 1. Detection of common runs of homozygosity according to different number of consecutive *SNPs*, percentage of animals, gap size and number of heterozygous genotypes.

Apêndices

Table S1. Parameter estimates for the nonlinear regression models by chromosome

Table S2. Parameter estimates for the nonlinear regression model for 5 replicates of 1 million randomly sampled r^2 values from each autosomal chromosome.

LISTA DE FIGURAS

CAPÍTULO 2 - ASSESSMENT OF AUTOZYGOSITY IN NELLORE COWS (*Bos indicus*) THROUGH HIGH-DENSITY SNP GENOTYPES

Figure 1. Frequency distribution of percentage of autosomal genome coverage by runs of homozygosity (F_{ROH}) of different minimum lengths (>0.5, >1, >2, >4, >8 and >16 Mb) and diagonal elements of the realized genomic relationship matrix ($G_{i,i}$).

Figure 2. Scatterplots (lower panel) and correlations (upper panel) of percentage of autosomal genome coverage by runs of homozygosity (F_{ROH}) of different minimum lengths (>0.5, >1, >2, >4, >8 and >16 Mb) and diagonal elements of the realized genomic relationship matrix ($G_{i,i}$).

Figure 3. Barplot of average percentage of chromosome coverage by runs of homozygosity (F_{ROH}) of minimum length of 0.5 Mb.

Figure 4. Manhattan plot of genome-wide locus autozygosity in Nellore cows. The dashed line represents the 50% threshold.

Apêndices

Figure S1. Frequency distribution of percentage of the total genome (autosomes + X) coverage by runs of homozygosity (F_{ROH}) of different minimum lengths (>0.5, >1, >2, >4, >8 and >16 Mb) and diagonal elements of the realized genomic relationship matrix ($G_{i,i}$).

Figure S2. Scatterplots (lower panel) and correlations (upper panel) of percentage of total genome (autosomes + X) coverage by runs of homozygosity (F_{ROH}) of different minimum lengths (>0.5, >1, >2, >4, >8 and >16 Mb) and diagonal elements of the realized genomic relationship matrix ($G_{i,i}$).

CAPÍTULO 1 – A AUTOZIGOZIDADE GENÔMICA E SEU POTENCIAL NO MELHORAMENTO GENÉTICO EM BOVINOS NO BRASIL

1. INTRODUÇÃO

Ao longo da história da evolução dos bovinos, vários foram os fenômenos que influenciaram a adaptação desses animais às distintas condições de sobrevivência, alguns dos quais inclusive deixaram marcas indelévels em seus genomas. Dentre esses fenômenos, destacam-se: i. Processos de seleção natural (ocorridos ao longo de milhares de anos) e artificial (com o acasalamento de animais selecionados pelo Homem), ii. Deriva genética ou aleatória (em que pequena parte da população passa a apresentar frequência alélica alterada); iii. Efeito de gargalo populacional (redução drástica da população e perda da variabilidade genética) e iv. Endogamia (como fruto do acasalamento entre indivíduos aparentados).

A permanência de alelos idênticos por descendência (*IBD*) no genoma dos bovinos ao longo da história evolutiva resultou no aumento da homozigosidade de determinados loci ao longo do mesmo. A autozigosidade, que pode ser determinada pela medida da ocorrência de segmentos genômicos homozigotos ao longo dos cromossomos, também denominados de corridas de homozigosidade (*ROH*), é resultado da manifestação combinada dos fenômenos acima descritos. Esse aumento da homozigosidade tanto pode se manifestar pela fixação de alelos desfavoráveis ou mutações, acarretando consequências negativas sobre as características de importância econômicas, quanto através da promoção de alelos favoráveis, levando ao aumento dos índices produtivos.

Com o intuito de detectar essas marcas no genoma, diversas tecnologias foram desenvolvidas ao longo das últimas duas décadas. Primeiramente, concluiu-se o sequenciamento e montagem do genoma bovino de referência, que por sua vez permitiu o desenvolvimento de testes analíticos para a detecção de milhares de marcadores moleculares do tipo *SNP* (*SNP chips*). Desta forma, tornou-se possível a estimação dos

níveis de endogamia genômicos nos bovinos, sendo o coeficiente de endogamia baseado nas corridas de homozigosidade (F_{ROH}), uma das alternativas de cálculo. Essa alternativa de cálculo consiste de trechos contínuos de segmentos autozigotos em indivíduos que possuem identidade por descendência (*IBD*). A identificação desses trechos, ou fragmentos homozigotos contínuos, ao longo do genoma dos bovinos, pode revelar a história da endogamia de uma população basicamente por dois aspectos: i. A identificação de fragmentos longos (i.e. > 4 Mb), que apontam para a ocorrência de endogamia/autozigosidade recente e ii. A observação de fragmentos curtos (i.e. < 4 Mb), que denotam a presença de endogamia ocorrida em tempos remotos.

Os padrões de *ROH*, além de discorrerem sobre a história da endogamia numa população, permitem a caracterização de alterações negativas ou positivas causadas por eventos evolutivos que influenciaram a adaptação dos bovinos desde tempos remotos. Consequentemente, essa nova informação pode permitir intervenções em programas de melhoramento genético animal, com o objetivo de selecionar e direcionar os acasalamentos, visando a promoção de características produtivas e rentáveis através da manipulação das *ROHs*.

2. REVISÃO DE LITERATURA

O termo endogamia, conhecido também como consanguinidade, refere-se ao efeito do cruzamento entre os indivíduos mais aparentados do que a média de uma determinada população, gerando descendentes que serão os pais da próxima geração (PEARL, 1913). Portanto, esse conceito aplica-se ao resultado do acasalamento entre indivíduos geneticamente relacionados, que apresentam um ou mais ancestrais comuns.

O cruzamento entre indivíduos aparentados possui importante papel tanto no melhoramento genético dos animais domésticos, quanto no decurso dos processos genéticos evolutivos (CHARLESWORTH & CHARLESWORTH, 1987, KRISTENSEN AND SORENSEN, 2005; CURIK et al., 2014).

Indivíduos endogâmicos possuem maior probabilidade de apresentar dois alelos idênticos por descendência num determinado locus, provocando o aumento da proporção de loci homozigotos no indivíduo e, eventualmente, na população a qual ele pertence

(DAVIS AND SIMMEN, 2010). A endogamia, portanto, é um fator de grande influência para a seleção de animais, pois fixa positivamente ou negativamente determinadas características através do aumento da homozigose.

Nesse contexto, a endogamia pode apresentar como consequências negativas, a diminuição da variabilidade genética e a ocorrência de danos à saúde genética da população, pois um rebanho com baixa variabilidade genética apresenta dificuldade de adaptação a mudanças ambientais (sejam elas naturais ou artificiais). Com isso, a capacidade de resposta a mudanças de um rebanho em condição de elevada endogamia será pequena, podendo levar ao que se denomina “depressão endogâmica” ou “por endogamia” (FERENČAKOVIĆ et al., 2012). Esse fenômeno pode afetar negativamente a expressão de uma característica produtiva ou reprodutiva do indivíduo, diminuindo seu desempenho ou sua capacidade de resposta a desafios (WRIGHT, 1922, KÖCK et al., 2009). A depressão endogâmica ocorre pelo aumento da homozigose (de loci específicos ou de todo o genoma) e consequente diminuição da variabilidade genética, sendo observada em animais, plantas e seres humanos (NISWANDER AND CHUNG, 1965, SCHULL et al., 1965; CHARLESWORTH & CHARLESWORTH, 1987).

A autozigosidade pode ocorrer por meio do acasalamento entre indivíduos aparentados ou pelo fenômeno de deriva genética (também conhecido como deriva aleatória), os quais aumentam a frequência de alelos homozigotos e a presença de variantes do tipo deletérias (geralmente recessivas) que normalmente se manifestam em indivíduos homozigotos. Dessa forma, para que ocorra a depressão endogâmica numa determinada característica, é necessária a presença de efeitos de dominância parcial, completa ou de sobredominância, tendo como principais consequências: i. A diminuição da variabilidade genética, ii. Perda de variantes presentes no DNA desses animais e iii. Alteração no desempenho populacional médio para características de interesse econômico (FALCONER & MACKAY, 1996).

Os efeitos da endogamia não são apenas aqueles de vertente negativa. Segundo Wright (1922), um dos efeitos positivos da endogamia pode ser o aumento da uniformidade do rebanho, que estaria correlacionado com o aumento da prepotência. Neste contexto, a prepotência em bovinos, isto é, a capacidade de um animal gerar progênie mais

uniformes, está relacionada com a produção de gametas também mais uniformes. Logo, animais aparentados e mais homozigotos do que a média da população, poderão ter menor diversidade na produção de gametas, resultando na uniformização do rebanho. O acasalamento dirigido desses animais homozigotos proporcionará progresso genético mais rápido do que aquele observado em acasalamentos aleatórios (CARVALHEIRO et al., 2004; NEVES et al., 2009).

A história do cálculo da endogamia teve início com Wright (1922), que definiu o coeficiente de endogamia como sendo a correlação entre os alelos advindos dos gametas reprodutivos femininos e masculinos em um novo ser. O nível de endogamia de um indivíduo foi determinado com base na estatística do coeficiente de endogamia (F) que, quando calculado a partir de dados completos de registros de pedigree, foram denominados de coeficiente de endogamia baseado no pedigree (F_{PED}).

Segundo Hall et al. (2012), o coeficiente de endogamia pode ser definido como a probabilidade de que dois alelos de um locus autossômico qualquer de um indivíduo sejam cópias de um único alelo presente em um de seus ancestrais. Ferencaković et al. (2012) definiram o coeficiente de endogamia como sendo a proporção do genoma de um indivíduo que é autozigoto, isto é, que possui alelos homozigotos idênticos por descendência (IBD). Segundo Crow (1954), um locus idêntico por descendência (IBD) é aquele que possui dois alelos herdados do mesmo haplótipo de um ancestral comum, tanto do pai quanto da mãe, na ausência de recombinação ou mutação (CURIK et al., 2014).

Os efeitos da endogamia sobre as características de interesse econômico foram estudados em várias raças bovinas existentes no Brasil tais como: Nelore (SANTANA et al., 2010), Guzerá (OLIVEIRA et al., 1999), Hereford (PARIACOTE et al., 1998), Simental (MAXIMINI et al., 2011), Gir (QUEIROZ et al., 1993), Marchigiana e Bonsmara (SANTANA et al., 2012). A **Tabela 1** sumariza os fenótipos analisados nos distintos estudos sobre o efeito da endogamia em raças bovinas no Brasil.

Tabela 1. Fenótipos (crescimento, reprodução e produção de leite) analisados em distintas raças bovinas no Brasil, segundo distintos autores.

Característica	Referência	Raça
Crescimento		
Peso a desmama	Santana et al., 2010	Nelore
Ganho de peso a desmama aos 18 meses de idade		
Escore visual para terminação (precocidade) aos 18 meses de idade		
Escore visual para musculatura aos 18 meses		
Altura da garupa		
Reprodução		
Circunferência escrotal aos 18 meses de idade		
Probabilidade de prenhes em novilhas aos 14 meses de idade		
Longevidade		
Crescimento		
Peso ao nascimento	Oliveira et al., 1999	Guzerá
Pesos aos 8 meses de idade		
Pesos aos 12 meses de idade		
Pesos aos 18 meses de idade		
Pesos aos 24 meses de idade		
Crescimento		
Peso ao nascimento	Pariacote et al., 1998	Hereford
Ganho de peso pré-desmama diário		
Peso a desmama		
Reprodução		
Volume	Maximini et al., 2011	Simental
Concentração		
Motilidade		
Número de espermatozóides por ejaculado		
Viabilidade dos espermatozoides		
Produção de Leite		
Produção do leite	Queiroz et al., 1993	Gir

Reprodução

Duração da gestação

Idade a primeira cria

Queiroz et al., 1993

Gir

Intervalos de partos

Crescimento

Peso ao nascimento aos 14 meses de idade

Santana et al.,

Marchigiana

Peso a desmama aos 205 dias

2012

Crescimento

Peso ao nascimento

Santana et al.,

Bonsmara

Peso a desmama

2012

Circunferência escrotal aos 12 meses de idade

Dentre as distintas características associadas ao crescimento avaliadas, diversas apresentaram decréscimo na média esperada dos indivíduos com distintos graus de endogamia, tais como: peso a desmama em animais da raça Nelore, Marchigiana, Bonsmara e Hereford, e peso ao nascimento nas raças Bonsmara e Hereford. O efeito negativo do aumento de 1% no coeficiente de endogamia sobre a característica peso ao nascimento foi observado nas raças Bonsmara e Hereford. Em contrapartida, nenhuma evidência de depressão endogâmica (DE) para a característica de peso ao nascimento foi relatada nas raças Marchigiana e Guzerá (PARIACOTE et al., 1998; OLIVEIRA et al., 1999; SANTANA et al., 2010; SANTANA et al., 2012).

Santana et al. (2010) analisaram o efeito da endogamia sobre as características de ganho de peso a desmama aos 18 meses de idade, escore visual para terminação (precocidade) aos 18 meses de idade, escore visual para musculatura aos 18 meses e altura da garupa na raça Nelore, descrevendo efeitos negativos sobre as mesmas, exceto para o escore visual para musculatura aos 18 meses.

Ainda no âmbito do crescimento, Oliveira et al. (1999) estudaram os efeitos da endogamia sobre as características peso ao nascimento (PN) e peso aos 8 (P8), 12 (P12), 18 (P18) e 24 (P24) meses de idade na raça Guzerá, evidenciando coeficientes de

regressão linear da endogamia negativos e significativos para P8, P12 e P24 de fêmeas e P18 de machos. Pariacote et al. (1998) observaram redução do ganho de peso pré-desmama diário na raça Hereford como efeito da depressão endogâmica.

Além dos efeitos da endogamia sobre características de crescimento, alguns poucos estudos foram conduzidos na esfera reprodutiva, os quais abordaram a circunferência escrotal aos 12 e 18 meses de idade nas raças Bonsmara e Nelore, apontando para a redução de - 0,126 cm e - 1,638 cm, respectivamente, quando ocorria 1% de aumento no coeficiente de endogamia (SANTANA et al., 2010; SANTANA et al., 2012).

A depressão endogâmica também foi relatada para outras características reprodutivas, tais como: probabilidade de prenhes em novilhas aos 14 meses e longevidade na raça Nelore (SANTANA et al., 2010), além da duração da gestação na raça Gir (QUEIROZ et al., 1993).

Características relacionadas com a qualidade do sêmen (volume, concentração, motilidade, número de espermatozóides por ejaculado e a viabilidade dos espermatozóides) foram analisadas sob a ótica do efeito da endogamia na raça Simental (Maximini et al., 2011). Apesar da baixa média dos coeficientes de endogamia (1,3%), todas as características de qualidade do sêmen apresentaram evidências de depressão endogâmica. Os maiores valores dos efeitos negativos da endogamia foram encontrados para as características viabilidade espermática ($-3,547 \pm 2,578 \%$) e número de espermatozoides ($-2,397 \pm 0,796 \times 10^9$). Em contrapartida, nenhuma evidência de DE foi identificada para as características de idade a primeira cria (IPC) e intervalos de partos (IEP) na raça Gir (QUEIROZ et al., 1993).

Os efeitos da endogamia sobre a produção de leite (PL) na raça Gir foram estudados por Queiroz et al. (1993) que observaram redução de produção da ordem de 19.99 Kg durante a lactação, quando ocorria aumento de 1% no coeficiente de endogamia.

Com base na literatura científica originada de estudos com populações brasileiras de bovinos, a depressão endogâmica afetou parcialmente distintas características associadas ao crescimento, reprodução e produção de leite, fatores cruciais para o bom desempenho e produção dos bovinos.

A depressão endogâmica causada pelos efeitos de alelos desfavoráveis de grande magnitude pode ser efetivamente reduzida através da seleção artificial monitorada por marcadores moleculares ou mesmo da deriva aleatória (CHARLESWORTH & CHARLESWORTH, 1987; HEDRICK, 1994; GLÉMIN, 2003).

No Brasil, a partir de meados da década de 30, foi criado o primeiro sistema de identificação de animais zebuínos (Registro Genealógico) (ABCZ, 2011), com o intuito de conferir maior precisão aos pedigrees das raças zebuínas. Dessa forma, o cálculo dos coeficientes de endogamia (F) passou a contar com informações mais precisas contidas nos registros genealógicos, mas que, entretanto, seguem sendo incompletas e contendo alta taxa de erro, o que pode levar ao cálculo de estimativas imprecisas, limitando assim o seu uso (RON et al., 1996; FERENČAKOVIĆ et al., 2012).

Graças a finalização do genoma bovino de referência (THE BOVINE GENOME SEQUENCING AND ANALYSIS CONSORTIUM et al., 2009; ZIMIN et al., 2009), a descoberta de variantes do tipo *SNP* (polimorfismo de sítio único, do inglês: *Single Nucleotide Polymorphism*) abundantes e comuns entre diversas raças bovinas (VAN TASSELL et al., 2008; THE BOVINE HAPMAP CONSORTIUM et al., 2009), e a disponibilidade comercial de microarranjos de DNA capazes de interrogar milhares de *SNP* ao longo de todo o genoma (MATUKUMALLI et al., 2009), tem sido possível a aplicação da genômica em programas de melhoramento genético de animais e plantas em larga escala e de forma rotineira (BJELLAND et al., 2003; AJMONE-MARSAN et al., 2010; LUAN et al., 2014).

Além da utilização das informações de pedigree para os cálculos do coeficiente de endogamia de um indivíduo ou de uma população homozigotos (isto é, que possuem dois alelos idênticos por estado (*IBS*), do inglês: *Identity by State*), agora é possível a estimação do coeficiente de endogamia (F) mais preciso a partir da informação genômica dos indivíduos (CURIK et al., 2014), através da detecção de segmentos homozigotos idênticos por descendência (*IBD*), do inglês: *Identity by Descent*, e que possuem um ancestral comum no genoma de um indivíduo (BROMAN & WEBER, 1999).

Um dos métodos mais empregados para a estimativa do coeficiente de endogamia é a corrida de homozigosidade (*ROH*), do inglês: *Runs of Homozygosity*, que detecta

segmentos autozigotos através dos padrões de marcadores moleculares do tipo *SNP* (KIM et al., 2013).

De acordo com Keller et al. (2011) o coeficiente de endogamia baseado na *ROH* (F_{ROH}), é suficientemente robusto para a estimativa da autozigosidade genômica de um indivíduo. Esse coeficiente é expresso como a proporção do genoma disposto em longos segmentos em homozigose, e apresenta alta correlação com os níveis de autozigosidade calculados por meio da genealogia (F_{PED}). Por basear-se na informação genotípica real do animal, ao invés de valores esperados por genealogia, o F_{ROH} apresenta-se como uma alternativa mais segura para a estimativa de níveis de endogamia, pois erros de genealogia em bovinos são muito comuns devido a erros de identificação e registros incompletos (RON et al., 1996; FERENČAKOVIĆ et al., 2011). Segundo Leutenegger et al. (2003) o (F_{ROH}) é definido como o comprimento do genoma oriundo da análise *ROH*, dividido pelo comprimento total do genoma coberto por marcadores moleculares do tipo *SNP*.

A *ROH* também pode ser definida como um trecho contínuo e ininterrupto da sequência de DNA no estado diplóide, com ausência de heterozigosidade (GIBSON et al., 2006; FERENČAKOVIĆ et al., 2013b), que é transmitido a progênie na forma de haplótipos idênticos (SÖLKNER et al., 2014). O comprimento dessas *ROH* no genoma pode indicar se a presença da endogamia na população provém de tempos remotos ou recentes.

Dessa forma, fragmentos longos em homozigose devem denotar autozigosidade recente, uma vez que a ocorrência de poucos eventos de recombinação mantém íntegros os haplótipos idênticos herdados a partir de um ancestral comum recente. Em contraste, fragmentos pequenos podem ou não resultar de autozigosidade proveniente de ancestrais remotos (CURIK et al., 2014; SÖLKNER et al., 2014). Por esse motivo, estimativas de autozigosidade levando em consideração fragmentos homozigotos muito pequenos podem apresentar viés, por inclusão de fragmentos homozigotos cujos alelos são idênticos apenas por estado e não por descendência. Em contraste, estimativas baseadas somente em fragmentos longos são limitadas à captura de informação de endogamia recente (FERENČAKOVIĆ et al., 2012). Dessa forma, pode-se considerar que o número e tamanho ideais para a análise *ROH* dependem de parâmetros e limiares impostos durante as etapas da análise (HOWRIGAN et al., 2011).

As regiões comuns de autozigosidade presentes no genoma de um indivíduo são denominadas de “ilhas de homozigosidade”. Segundo Nothnagel et al. (2010), estas regiões em seres humanos, quando se encontram presentes em mais de 50% dos indivíduos de uma população, poderiam indicar a ocorrência de forte seleção. Os padrões de *ROH*, que em última instância refletem a autozigosidade, podem ser resultado de vários fenômenos ao longo da história da evolução, tais como: deriva genética, gargalos populacionais, endogamia e a seleção natural ou artificial (KELLER et al., 2011; PEMBERTON et al., 2012; PURFIELD et al., 2012),

Desde tempos remotos, grandes filósofos e naturalistas discutiram sobre a adaptação de seres vivos ao ambiente, o que gerou várias teorias ao longo do tempo. Darwin foi um grande observador da natureza, que abordou a história da biologia evolutiva que, segundo ele, era explicada pela adaptação dos seres vivos baseada na seleção natural. O termo adaptação refere-se ao comportamento natural de um indivíduo que impacta na evolução da espécie, ou mais especificamente a presença de uma determinada característica que confere a esse indivíduo a capacidade de se adaptar, sobreviver e reproduzir em um ambiente extremamente diferente daquele de seu nascimento (DARWIN, 1859).

Ao longo das gerações, as raças animais passam pelo processo de seleção natural, que modifica seus genótipos e, conseqüentemente, seus fenótipos, garantindo sua sobrevivência e gerando animais adaptados e resistentes a determinadas doenças (PORTO-NETO et al., 2012). Raças também podem ser extintas devido a processos de seleção, nos quais a substituição de populações inteiras ou o cruzamento indiscriminado de animais selecionados de raças específicas causam impacto na composição genética dos rebanhos (AJMONE-MARSAN et al., 2010).

Neste contexto, os bovinos foram fortemente expostos ao processo de seleção artificial ao longo das gerações, o que deve ter contribuído para deixar marcas nos genomas desses animais. Segundo Zhang et al. (2014), regiões genômicas sob forte pressão de seleção apresentaram níveis elevados de *ROH*. O aumento das *ROH* nos genomas animais pode levar a efeitos fenotípicos prejudiciais, devido ao acúmulo de mutações advindas dos sistemas de acasalamento com preponderância de pequenos

grupos de animais (GLÉMIN, 2003), os quais podem dificultar o processo de adaptação. De acordo com Hedrick (1994), uma forma de evitar esses efeitos negativos é a condução de cruzamentos entre parentes realizados de forma programada e selecionada, a fim de manter a diversidade genética e permitir o processo de adaptação desses animais futuramente.

Outro fenômeno que pode ter contribuído para o aumento da autozigosidade ao longo da história dos bovinos foi a deriva genética ou aleatória. Este fenômeno consiste de evento evolutivo no qual uma população pequena de animais sofre alteração estocástica no genoma, que se nota através da alteração das frequências alélicas esperadas (SALMAN, 2007). Este tipo evento evolutivo foi importante para a formação das raças bovinas, que se fundamentam invariavelmente no isolamento geográfico e adaptação a nichos ecológicos diferentes, acumulando diferenças alélicas ao longo do tempo (EGITO et al., 2002).

Em adição, o gargalo populacional também pode contribuir para a ocorrência de autozigosidade, e pode ser devido a extinção, ou impossibilidade de reprodução, de parte significativa da população, resultando na diminuição dos possíveis reprodutores. No melhoramento genético, o gargalo populacional pode ser consequência do uso massivo de touros em programas de inseminação artificial (FARIA et al., 2002; FILHO et al., 2002).

Além dos fenômenos evolutivos descritos acima, a domesticação dos bovinos zebuínos (*Bos indicus*) e taurinos (*Bos taurus*) que precede o contato do Homem com os bovinos selvagens ou *aurochs* (*Bos primigenius*) (ZEUNER, 1963), também pode ter contribuído com o aumento da autozigosidade genômica ao longo dos anos, devido a seleção e aos acasalamentos dirigidos realizados nesses animais.

Todos os eventos evolutivos, somados ao processo de domesticação e seleção dos bovinos resultaram na pecuária de corte moderna. A pecuária de corte brasileira, em especial a que compreende a raça Nelore, vive atualmente um momento particular no qual a seleção genética intensiva para características de importância econômica visam a produção de animais geneticamente superiores e homogêneos para determinados fenótipos. Pode-se inferir, portanto que, o aumento da ocorrência de alelos autozigotos no genoma dos bovinos zebuínos (*Bos indicus*) e taurinos (*Bos taurus*), acumulados ao longo

dos anos (em passado recente ou remoto), foi causado por eventos evolutivos (tais como a endogamia, seleção natural e artificial, deriva genética ou aleatória e gargalo populacional) descritos nessa revisão, podem ser compreendidos e explorados graças a disponibilidade das poderosas ferramentas analíticas disponíveis na atualidade.

3. REFERÊNCIAS

ABCZ - Associação Brasileira dos Criadores de Zebu. Disponível em: <http://www.abcz.org.br/AreaTecnica/RegistroGenealogico/Historia>. História do Registro Genealógico. Acessado em: 15 de Fevereiro de 2011. 2011.

AJMONE-MARSAN, P.; GARCIA, J. F.; LENSTRA, J. A. AND THE GLOBALDIV CONSORTIUM. On the Origin of Cattle: How Aurochs Became Cattle and Colonized the World. *Evolutionary Anthropology*. 19, 148-157, 2010.

BJELLAND, D. W.; WEIGEL, K. A.; VUKASINOVIC, N.; NKURUMAH, J. D. Evaluation of inbreeding depression in Holstein cattle using whole-genome *SNP* markers and alternative measures of genomic inbreeding. *J. Dairy Sci.* 96(7), 4697-4706, 2013.

BROMAN K.W., WEBER J.L. Long homozygous chromosomal segments in reference families from the Centre d'E tude du Polymorphisme Humain. *Am. J. Hum. Genet.* 65, 1493-1500, 1999.

CARVALHEIRO, R.; PIMENTEL, E. DA C. G. Endogamia: possíveis consequências e formas de controle em Programas de Melhoramento de bovinos de corte. *Anais. II GEMPEC - Workshop em genética e melhoramento na pecuária de corte*, 2004.

CHARLESWORTH, D. & CHARLESWORTH, B. Inbreeding depression and its evolutionary consequences. *Annu. Rev. Ecol. Syst.* 18, 237-268, 1987.

CROW, J. F. Breeding structure of populations. II. Effective population number. In: KEMPTHORN, E., O.; BANCROFT, T. A.; GOWEN, J. W.; LUSH, J. L. (Eds.), *Statistics and Mathematics in Biology*. Iowa State University PRESS, AMES, USA, 543-556, 1954.

CURIK, I.; FERENČAKOVIĆ, M.; SÖLKNER, J. Inbreeding and runs of homozygosity: A possible solution to an old problem. *Livestock Science*. 166, 26-34, 2014.

DARWIN, C. "A Origem das Espécies" (On the Origin of Species by Means of Natural Selection, on the Preservation of Favoured Races in the Struggle for Life (Sobre a Origem das Espécies por Meio da Seleção Natural ou a Preservação de Raças Favorecidas na Luta pela Vida). London: John Murray, Albemarle street, p.502, 1859.

DAVIS, M. E. AND SIMMEN, R. C. M. Estimates of inbreeding depression for serum insulin-like growth factor I concentrations, body weights, and body weight gains in Angus beef cattle divergently selected for serum insulin-like growth factor I concentration. *J. Anim. Sci.* 88, 552-561, 2010.

EGITO, A. A.; MARIANTE, A. S.; E ALBUQUERQUE, M. S. M. Programa Brasileiro de Conservação de Recursos Genéticos Animais. The Brazilian Genetic Resources Conservation Programm. *Arch. Zootec.* 51, 39-52, 2002.

FALCONER, D. S.; MACKAY, T. F. C. Introduction to Quantitative Genetics 4th ed. Essex, England: *Longman*. 1996.

FARIA, F. J. C.; FILHO, A. E. V.; MADALENA, F. E.; JOSAHKIAN, L. A. Estrutura populacional da raça Nelore Mocho. *Arq. Bras. Med. Vet. Zootec.* 54(5), 2002.

FERENČAKOVIĆ, M.; HAMZIC, E.; GREDLER, B.; CURIK, I.; SOLKNER, J. Runs of homozygosity reveal genome-wide autozygosity in the Austrian fleckvieh cattle. *Agric. Conspec. Sci.* 76, 325-328, 2011.

FERENČAKOVIĆ, M.; HAMZIC, E.; GREDLER, B.; SOLBERG, T. R.; KLEMETSDAL, G.; CURIK, I. & SOLKNER, J. Estimates of autozygosity derived from runs of homozygosity: empirical evidence from selected cattle populations. *Journal Animal Breeding and Genetics*, 1-8, 2012.

FERENČAKOVIĆ, M.; SÖLKNER, J.; CURIK, I. Estimating autozygosity from high-throughput information: effects of SNP density and genotyping errors. *Genet. Sel. Evol.* 45, 42, 2013b.

FILHO, A. E. V.; FARIA, F. J. C.; MADALENA, F. E. E JOSAHKIAN, L. A. Estrutura populacional do rebanho Indubrasil registrado no Brasil. *Arch. Latinoam. Prod. Anim.* 10(2), 86-92, 2002.

GIBSON, J.; MORTON, N. E.; COLLINS, A. Extended tracts of homozygosity in outbred human populations. *Hum. Mol. Genet.* 15, 789-795, 2006.

GLÉMIN, S. How are deleterious mutations purged? Drift versus nonrandom mating. *Evolution*. 57, 2678-2687, 2003.

HALL, N.; MERCER, L.; PHILLIPS, D.; SHAW, J. AND ANDERSON, A. D. Maximum likelihood estimation of individual inbreeding coefficients and null allele frequencies. *Genet. Res. Camb.* 94, 151-161, 2012.

HEDRICK, P.W. Purging inbreeding depression and the probability of extinction: full-sib mating. *Heredity*. 73, 363-372, 1994.

HOWRIGAN, D.; SIMONSON, M.; KELLER, M.: Detecting autozygosity through runs of homozygosity: A comparison of three autozygosity detection algorithms. *BMC Genomics*. 12(1):460, 2011.

KELLER, M. C.; VISSCHER, P. M.; GODDARD, M. E. Quantification of inbreeding due to distant ancestors and its detection using dense single nucleotide polymorphism data. *Genetics*. 89, 237-249, 2011.

KIM, EUI-SOO; COLE, J. B.; HUSON, H.; WIGGANS, G. R.; VAN TASSEL, C. P.; CROOKER, B. A.; LIU, G.; DA, Y.; SONSTEGARD, T. S. Effect of Artificial Selection on Runs of Homozygosity in U.S. Holstein Cattle. *PLOS One*. 8(11), e80813, 2013.

KÖCK A.; FURST-WALTL B.; BAUMUNG R. Effects of inbreeding on number of piglets born total, born alive and weaned in Austrian Large White and Landrace pigs. *Arch. Tierz.* 52, 51-64, 2009.

KRISTENSEN, T.; SORENSEN, A. Inbreeding-lessons from animal breeding, evolutionary biology and conservation genetics. *Anim. Sci.* 80, 121-133, 2005.

LEUTENEGGER, A.; PRUM, B.; GENIN, E.: Estimation of the inbreeding coefficient through use of genomic data. *Am J Hum Genet.* 73(516), 523, 2003.

LUAN, T.; YU, X.; DOLEZAL, M.; BAGNATO, A.; and MEUWISSEN, T. H. E. Genomic prediction based on runs of homozygosity. *Genet. Sel. Evol.* 46:64, 2014.

MATUKUMALLI, L. K.; LAWLEY, C. T.; SCHNABEL, R. D.; TAYLOR, J. F.; ALLAN, M. F.; HEATON, M. P.; O'CONNELL, J.; MOORE, S. S.; SMITH, T. P. L.; SONSTEGARD, T. S.; VAN TASSEL, C. P. Development and Characterization of a High Density *SNP* Genotyping Assay for Cattle. *PLOS ONE*. 4(4), e5350, 2009.

MAXIMINI, L.; FUERST-WALTL, B.; GREDLER, B. AND BAUMUNG, R. Inbreeding Depression on Semen Quality in Australian Dual – Purpose Simmental Bulls. Short Communication. *Reproduction in Domestic Animals*. 46, e102-e104, 2011.

NEVES, H. H. de R.; CARVALHEIRO, R.; CARDOSO, V.; FRIES, L. A.; DE QUEIROZ, S. A. Acasalamento dirigido para aumentar a produção de animais geneticamente superiores e reduzir a variabilidade da progênie em bovinos. *R. Bras. Zootec.* 38(7), 1201-1204, 2009.

NISWANDER, J. D. AND CHUNG, C. S. The Effects of Inbreeding on Tooth Size in Japanese Children. *American Journal of Human Genetics*. 17(5), 1965.

NOTHNAGEL M., T. LU, M. KAYSER, M. KRAWCZAK. Genomic and geographic distribution of *SNP*-defined runs of homozygosity in Europeans. *Human Molecular Genetics*. 19, 2927 -2935, 2009.

OLIVEIRA, J. A. de; BASTOS, J. F. P.; TONHATI, H. Endogamia em um Rebanho da Raça Guzerá. *Rev. Bras. Zootec.* 28(4), 721-728, 1999.

PARIACOTE, F.; VAN VLECK, L. D. and MACNEIL, M. D. Effects of inbreeding and heterozygosity on preweaning traits in a closed population of Herefords under selection. *ANIM SCI*. 76:1303-1310, 1998.

PEARL, R. A contribution towards an analysis of the problem of inbreeding. *The American Naturalist*. v. XLVII (562). 1913.

PEMBERTON , T., ABSHER, D., FELDMAN , M., MYERS, R., ROSENBERG, N., LI, J. Genomic patterns of homozygosity in world wide human populations. *Am. J. Hum. Genet.* 91, 275-292, 2012.

PORTO-NETO, L. R.; SONSTEGARD, T. S.; LIU, G. E.; BICKHART, D. M.; DA SILVA, M. V. B.; MACHADO, M. A.; UTSUNOMIYA, Y. T.; GARCIA, J. F.; GONDRO, C.; AND VAN TASSELL, C. P. Genomic divergence of zebu and taurine cattle identified through high-density *SNP* genotyping. *BMC Genomics.* 14, 876. <http://www.biomedcentral.com/1471-2164/14/876>, 2013.

PURFIELD, D.; BERRY, D.; MCPARLAND, S.; BRADLEY, D. Runs of homozygosity and population history in cattle. *BMC Genetics*, 13 (1): 70, 2012.

QUEIROZ, S. A. de; LÔBO, R. B.; MARTINEZ, M. L. Efeito da endogamia sobre algumas características de importância econômica na raça *Gir*. *Revista Sociedade Brasileira de Zootecnia.* 22(5), 773-786, 1993.

RON, M.; BLANC, Y.; BAND, M.; EZRA, E.; WELLER, J. I. Misidentification rate in the Israeli dairy cattle population and its implications for genetic improvement. *J. Dairy Sci.* 79, 676-681, 1996.

SALMAN, A. K. D. Conceitos básicos de genética de Populações. Porto Velho, RO: Embrapa Rondônia. 27p. : il. color. (Documentos / Embrapa Rondônia, ISSN 0103-9865;118).1. Genética animal. 2. Melhoramento genético. I. Título. II. Série.CDD (21.ed.) 636.082, 2007.

SANTANA JR., M. L.; OLIVEIRA, P. S.; PEDROSA, V. B.; ELER, J. P.; GROENEVELD, E.; FERRAZ, J. B. S. Effect of inbreeding on growth and reproductive traits of Nellore cattle in Brazil. *Livestock Science.* 131, 212-217, 2010.

SANTANA JR., M. L.; OLIVEIRA, P. S.; ELER, J. P.; GUTIÉRREZ, J. P. AND FERRAZ, J. B. S. Pedigree analysis and inbreeding depression on growth traits in Brazilian Marchigiana and Bonsmara breeds. *J ANIM SCI.* 90:99-108, 2012.

SCHULL, W. J.; NEEL, J. V. The Effect of Inbreeding on Japanese Children. *Harper & Row*, New York, USA. 1965.

SÖLKNER, J.; FERENČAKOVIĆ, M.; KARIMI, Z.; PEREZ O'BRIEN, A.M.; MÉSZÁROS, G.; EAGLEN, S.; BOISON, S.A.; AND CURIK, I. Extremely Non-uniform: Patterns of Runs of Homozygosity in Bovine Populations. *Proceedings, 10th World Congress of Genetics Applied to Livestock Production*. 2014.

THE BOVINE GENOME SEQUENCING AND ANALYSIS CONSORTIUM; ELSIK, C. G.; TELLAM, R. L.; WORLEY, K. C. The genome sequence of taurine cattle: a window to ruminant biology and evolution. *Science*. 324, 522-528, 2009.

THE BOVINE HAPMAP CONSORTIUM. Genome-Wide survey of *SNP* variation uncovers the genetic structure of cattle breeds. *Science*. 324, 528-532, 2009.

VAN TASSELL, C. P.; SMITH, T. P.; MATUKUMALLI, L. K.; TAYLOR, J. F.; SCHNABEL, R. D.; LAWLEY, C. T.; HAUDENSCHILD, C. D.; MOORE, S. S.; WARREN, W. C.; SONSTEGARD, T. S. *SNP* discovery and allele frequency estimation by deep sequencing of reduced representation libraries. *Nature Methods*. 5(3), 247-252, 2008.

WRIGHT, S. Coefficients of inbreeding and relationship. *Am. Nat.* 56, 330-338, 1922.

ZEUNER, F. E. A history of domesticated animals. *London: Hutchinson*. 1963.

ZHANG, Q.; GULDBRANDTSEN, B.; BOSSE, M.; LUND, M. S.; SAHANA, G. Runs of homozygosity and distribution of functional variants in cattle genome. *Proceedings, 10th World Congress of Genetics Applied to Livestock Production*. 2014.

ZIMIN, A. V.; DELCHER, A. L.; FLOREA, L.; KELLEY, D. R.; SCHATZ, M. C.; PUIU, D.; HANRAHAN, F.; PERTEA, G.; VAN TASSELL, C. P.; SONSTEGARD, T. S.; MARÇAIS, G.; ROBERTS, M.; SUBRAMANIAN, P.; YORKE, J. A. AND SALZBERG, S. L. A whole-genome assembly of the domestic cow, *Bos taurus*. *Genome Biology*. 2009.

CAPÍTULO 2 - ASSESSMENT OF AUTOZYGOSITY IN NELLORE COWS (*Bos indicus*) THROUGH HIGH-DENSITY SNP GENOTYPES

Ludmilla B. Zavarez¹, Yuri T. Utsunomiya¹, Adriana S. Carmo¹, Haroldo H.R. Neves^{2,3}, Roberto Carneiro², Maja Ferencaković⁴, Ana M. Pérez O'Brien⁵, Ino Curik⁴, John B. Cole⁶, Curtis P. Van Tassell⁶, Marcos V.G.B. da Silva⁷, Tad S. Sonstegard⁶, Johann Sölkner⁵, José F. Garcia^{1,8*}

Frontiers in Genetics (Accepted for publication)

¹UNESP – Univ. Estadual Paulista, Faculdade de Ciências Agrárias e Veterinárias, Departamento de Medicina Veterinária Preventiva e Reprodução Animal, Jaboticabal, São Paulo, Brazil

²UNESP – Univ. Estadual Paulista, Faculdade de Ciências Agrárias e Veterinárias, Departamento de Zootecnia, Jaboticabal, São Paulo, Brazil

³GenSys Consultores Associados, Porto Alegre, Rio Grande do Sul, Brasil

⁴Department of Animal Science, Faculty of Agriculture, University of Zagreb, Zagreb, Croatia

⁵BOKU - University of Natural Resources and Life Sciences, Department of Sustainable Agricultural Systems, Division of Livestock Sciences, Vienna, Austria

⁶Animal Genomics and Improvement Laboratory, United States Department of Agriculture, Agricultural Research Service, Beltsville, Maryland, United States of America

⁷Bioinformatics and Animal Genomics Laboratory, Embrapa Dairy Cattle, Juiz de Fora, Minas Gerais, Brazil

⁸Laboratório de Bioquímica e Biologia Molecular Animal, UNESP – Univ. Estadual Paulista, Faculdade de Medicina Veterinária de Araçatuba, Departamento de Apoio, Produção e Saúde Animal, Araçatuba, São Paulo, Brazil

***Correspondence:** José Fernando Garcia, Laboratório de Bioquímica e Biologia Molecular Animal, UNESP – Univ. Estadual Paulista, Faculdade de Ciências Agrárias e Veterinárias, Departamento de Apoio, Produção e Saúde Animal, Rua Clóvis Pestana 793, Araçatuba, São Paulo, 16050-680, Brazil.

jfgarcia@fmva.unesp.br

ABSTRACT – The use of relatively low numbers of sires in cattle breeding programs, particularly on those for carcass and weight traits in Nellore beef cattle (*Bos indicus*) in Brazil, has always raised concerns about inbreeding, which affect conservation of genetic resources and sustainability of this breed. Here, we investigated the distribution of autozygosity levels based on runs of homozygosity (*ROH*) in a sample of 1,278 Nellore cows, genotyped for over 777,000 *SNPs*. We found *ROH* segments larger than 10 Mb in over 70% of the samples, representing signatures most likely related to the recent massive use of few sires. However, the average genome coverage by *ROH* (> 1 Mb) was lower than previously reported for other cattle breeds (4.58%). In spite of 99.98% of the *SNPs* being included within a *ROH* in at least one individual, only 19.37% of the markers were encompassed by common *ROH*, suggesting that the ongoing selection for weight, carcass and reproductive traits in this population is too recent to have produced selection signatures in the form of *ROH*. Three short-range highly prevalent *ROH* autosomal hotspots (occurring in over 50% of the samples) were observed, indicating candidate regions most likely under selection since before the foundation of Brazilian Nellore cattle. The putative signatures of selection on chromosomes 4, 7 and 12 may be involved in resistance to infectious diseases and fertility, and should be subject of future investigation.

Keywords: *Bos indicus*₁, Runs of Homozygosity₂, Selection₃, Cattle₄, Fertility₅, Disease resistance₆.

1. Introduction

Autozygosity is the homozygote state of identical-by-descent alleles, which can result from several different phenomena such as genetic drift, population bottleneck, mating of close relatives, and natural and artificial selection (Falconer & MacKay, 1996; Keller et al., 2011; Curik et al., 2014). In the past 20 years, the heavy use of relatively low number of sires in Brazilian Nellore breeding programs (*Bos indicus*) is deemed to have mimicked all these triggers of autozygosity, especially considering the increasing use of artificial insemination over the decades. As inbreeding has been incriminated in reduced fitness and reproductive performance in other cattle populations under artificial selection (Bjelland et al., 2013; Leroy et al., 2014), avoidance of mating of close relatives is a typical practice of many Nellore breeders. Therefore, there is a growing interest in characterizing and monitoring autozygosity in this breed to preserve genetic diversity and allow the long-term sustainability of breeding programs in Brazil.

Recently, the use of high-density single nucleotide polymorphism (*SNP*) genotypes to scan individual genomes for contiguous homozygous chromosomal fragments has been proposed as a proxy for the identification of identical-by-descent haplotypes (Gibson et al., 2006; Lencz et al., 2007). As the length of autozygous chromosomal segments is proportional to the number of generations since the common ancestor (Howrigan et al., 2011), the identification of runs of homozygosity (*ROH*) can reveal recent and remote events of inbreeding, providing invaluable information about the genetic relationships and demographic history of domesticated cattle (Purfield et al., 2012; Ferenčaković et al., 2013a; Kim et al., 2013). Also, given the stochastic nature of recombination, the occurrence of *ROH* is highly heterogeneous across the genome, and hotspots of *ROH* across a large number of samples (here after referred as common *ROH*) may be indicative of selective pressure. Moreover, the fraction of an individual's genome covered by *ROH* can be used as an estimate of its genomic autozygosity or inbreeding coefficient (McQuillan et al., 2008; Curik et al., 2014).

Here, we investigated the occurrence of *ROH* in the genomes of 1,278 Nellore cows (*Bos indicus*) genotyped for over 777,000 *SNPs*. We aimed at

characterizing the distribution of *ROH* length and genome-wide levels of autozygosity, as well as detecting common *ROH* that may be implicated in past events of selection.

2. Materials and methods

2.1. Ethical statement

The present study was exempt of the local ethical committee evaluation as genomic DNA was extracted from stored hair samples of animals from commercial herds.

2.2. Genotyping and data filtering

A total of 1,278 cows were genotyped with the Illumina[®] BovineHD Genotyping BeadChip assay (HD), according to the manufacturer's protocol. These animals comprised part of the genomic selection reference population from a commercial breeding program. These dams were born between 1993 and 2008, being under routine genetic evaluation for weight, carcass and reproductive traits by the DeltaGen program, an alliance of Nellore cattle breeders from Brazil. Data filtering was performed using *PLINK* v1.07 (Purcell et al., 2007), and markers were removed from the dataset if GenTrain score lower than 70% or a call rate lower than 98% was observed. All genotyped samples exhibited call rates greater than 90%, thus no animals were filtered from further analyses. Minor allele frequency (MAF) was not used as an exclusion criterion in this analysis, so that the detection of homozygous segments was not compromised. Both autosomal and X-linked markers were included.

2.3. Estimates of genomic individual autozygosity

Genomic autozygosity was measured based on the percentage of an individual's genome that is covered by *ROH*. Stretches of consecutive homozygous genotypes were identified for each animal using *SNP & Variation Suite* v7.6.8 (Golden Helix, Bozeman, MT, USA <http://www.goldenhelix.com>), and chromosomal segments were declared *ROH* under the following criteria: 30 or more consecutive homozygous *SNPs*, a density of at least 1 *SNP* every 100 kb, gaps of no more than 500 kb between *SNPs*, and no more than 5 missing genotypes across all individuals. In order to account for genotyping error and avoid underestimation of long *ROH* (Ferenčaković et al., 2013b), heterozygous genotype calls were allowed under conditions where there were 2 heterozygous genotypes for $ROH \geq 4\text{Mb}$, or no heterozygous genotypes for $ROH < 4\text{Mb}$. Autozygosity was estimated according to McQuillan and colleagues (2008):

$$F_{ROH} = \frac{\sum_{j=1}^n L_{ROH_j}}{L_{total}}$$

Where L_{ROH_j} is the length of *ROH* j , and L_{total} is the total size of the genome covered by markers, calculated from the sum of inter-marker distances in the UMD v3.1 assembly. In order to facilitate comparisons with other studies, F_{ROH} was calculated using both the genome size based on autosomal and autosomal + X chromosomes. For each animal, F_{ROH} was calculated based on *ROH* of different minimum lengths: 0.5, 1, 2, 4, 8 or 16 Mb, representing autozygosity events that occurred approximately 100, 50, 25, 13, 6 and 3 generations in the past, respectively (Howrigan et al., 2011; Ferenčaković et al., 2013b). Additionally, chromosome-wise F_{ROH} was also computed.

An alternative measure of autozygosity was obtained by computing the diagonal elements of a modified realized genomic relationship matrix (VanRaden 2008; VanRaden et al., 2011), calculated as:

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

Where Z is a centered genotype matrix and p_l is the reference allele frequency at locus l . Matrix Z is obtained by subtracting from the genotype matrix M (with genotype scores coded as 0, 1 or 2 for alternative allele homozygote, heterozygote and reference allele homozygote, respectively) the matrix P , whose elements of column l are equal to $2p_l$. The diagonal elements of G ($G_{i,i}$) represent the relationship of an animal with itself, and thus encapsulate autozygosity information. Following VanRaden et al. (2011), $G_{i,i}$ can provide a more suitable proxy for the pedigree-based inbreeding coefficient when assuming $p_l = 0.5$, rather than using base population allele frequencies estimates (which could be difficult to estimate especially in absence of complete pedigree data). Thus, matrix G was computed using allele frequencies fixed at 0.5.

2.4. Detection of common runs of homozygosity

Chromosomal segments presenting *ROH* hotspots were defined as *ROH* islands or common *ROH*. In order to identify such genomic regions, we used two different strategies. First, we used the clustering algorithm implemented in *SNP & Variation Suite v7.6.8*, which identifies clusters of contiguous set of *SNPs* with size $>s_{min}$, where every *SNP* has at least n_{min} samples presenting a run. Clusters were identified based on a fixed minimum cluster size of $s_{min} = 0.5$ Mb for varying minimum number of samples: 127 (10%), 255 (20%), 319 (25%), and 639 (50%). In order to assess the sensitivity of the algorithm to parameter settings in *ROH* detection, we repeated the analysis using minimum numbers of 30 or 150 *SNPs* in a run, maximum gap sizes of 100 kb or 500 kb, and 0 or 2 heterozygous genotypes as variable parameters.

Alternatively, we calculated locus autozygosity (F_L) following Kim et al. (2013). Briefly, for each *SNP*, animals were classified as autozygous (1) or non-autozygous (0) based on the presence of a *ROH* encompassing the *SNP*. Then, F_L was simply computed as the percentage of animals with autozygosity status equal to 1.

3. Results and Discussion

The ultimate goal in livestock breeding strategies is to improve populations by selecting animals with above-average productive performance as parents of the next generation. Unavoidably, the use of ever-smaller numbers of animals as founders incurs in autozygosity. Evidence from whole-genome sequencing studies in humans indicate that highly deleterious variants are common across healthy individuals (MacArthur et al., 2012; Xue et al., 2012), and although no such systematical survey has been conducted in cattle to the present date, it is highly expected that unfavorable alleles also segregate in cattle populations. Therefore, inbreeding is expected to inadvertently increase autozygosity of such unfavorable alleles (Szpiech et al., 2013), potentially causing economic losses. Here, we analyzed high-density *SNP* genotypes in order to characterize autozygosity in the genomes of a sample of Nellore cows under artificial selection for weight, carcass and reproductive traits.

3.1. Distribution of *ROH* length

After filtering, 668,589 *SNP* marker genotypes across 1,278 animals were retained for analyses. The average, median, minimum and maximum *ROH* length detected across all chromosomes were 1.26, 0.70, 0.50 and 70.91 Mb, respectively, suggesting this specific Nellore cattle population experienced both recent and remote autozygosity events. Segments as large as 10 Mb are traceable to inbreeding that occurred within the last five generations (Howrigan et al., 2011), and a total of 942 samples (73.7%) presented at least one homozygous fragment larger than 10 Mb. Therefore, it is likely that these long *ROH* are signatures of the extended use of recent popular sires.

3.2. Distribution of genome-wide autozygosity

The distributions of $G_{i,i}$ and F_{ROH} based on autosomal *ROH* of different minimum lengths (>0.5, >1, >2, >4, >8 or >16 Mb) are shown in **Figure 1**.

Although the inclusion of the X chromosome did not cause substantial differences in the calculation of genome-wide F_{ROH} (**File S1 – Figure S1**), we focused on the estimates using only autosomes for the ease of comparison with other studies. The skewness of the autosomal F_{ROH} distribution increased as the minimum fragment length increased, ranging from 1.56 for $F_{ROH>0.5Mb}$ to 3.98 for $F_{ROH>16Mb}$. The number of animals with $F_{ROH} = 0$ also increased as the minimum ROH length increased, starting at 12 (0.94%) for $F_{ROH>2Mb}$ and increasing to 827 (64.71%) for $F_{ROH>16Mb}$. These findings suggest that varying the minimum ROH length in the calculation of F_{ROH} can be useful to discriminate animals with recent and remote autozygosity.

As shown in **Figure 2**, the correlation between autosomal $F_{ROH>1Mb}$ and $G_{i,i}$ ($r = 0.69$) was close to the ones reported by Ferenčaković et al. (2013b) for the comparison between $F_{ROH>1Mb}$ derived from the HD panel and pedigree estimates in Brown Swiss ($r = 0.61$), Pinzgauer ($r = 0.62$), and Tyrol Grey ($r = 0.75$). Similar correlations were observed when the X chromosome was included in the analysis (**File S1 - Figure S2**). McQuillan et al. (2008) also reported correlations between F_{ROH} and pedigree estimates in human European populations ranging from 0.74 to 0.82. Considering that VanRaden (2008) proposed G as a proxy for a numerator relationship matrix obtained from highly reliable and recursive pedigree data, we expect that the correlations found for $G_{i,i}$ are fair approximations to the ones we would have found if complete pedigree data was available.

In the present study, correlations between F_{ROH} and $G_{i,i}$ decreased as a function of different ROH length (**Figure 2**). This may be due to the properties of the G matrix, which is based on individual loci, whereas F_{ROH} is based on chromosomal segments. Ferenčaković et al. (2013b) showed that medium density *SNP* panels, such as the Illumina® BovineSNP50, systematically overestimate F_{ROH} when segments shorter than 4 Mb are included in the calculations, while the Illumina® BovineHD panel is robust for the detection of shorter segments. Hence, although the inclusion of short length ROH in the calculation of F_{ROH} may be desirable for autozygosity estimates accounting for remote inbreeding, there is a compromise between *SNP* density, minimum ROH length and false discovery of ROH . Since the HD panel allows for the detection of short ROH , in this section we focused on the results obtained with $F_{ROH>1Mb}$

as it presented the second highest correlation with $G_{i,i}$ and is comparable with previous studies.

The minimum, average, median, and maximum autosomal $F_{ROH>1Mb}$ across all animals were 0.43, 4.79, 4.58, and 18.55%, respectively. The animal presenting the highest autozygosity value (18.55%) exhibited 69 $ROH > 1$ Mb encompassing 465.66 Mb of the total autosomal genome extension covered by markers (2.51 Gb), with a mean ROH length of 6.75 ± 9.20 Mb, and a maximum segment length of 43.79 Mb. The least inbred animal presented 8 $ROH > 1$ Mb, summing up only 10.72 Mb, with an average length of 1.34 ± 0.46 Mb and a maximum of 2.43 Mb.

The coefficient of variation (here denoted as the ratio of the standard deviation to the mean) of the $F_{ROH>1Mb}$ distribution was 37.5%, indicating moderate variability in autozygosity levels in this sample. In spite of the average genome coverage by ROH of 4.58% may seem to indicate moderate inbreeding levels for classical standards, it has to be considered that incomplete pedigree data usually fails to capture remote inbreeding, so that traditional inbreeding estimates based on pedigree are only comparable with F_{ROH} calculated over large ROH lengths, which in the present study were close to 0%.

Compared to other cattle populations, this sample of Nellore cows presented a lower average autozygosity. For instance, Ferencaković et al. (2013b) reported average autosomal $F_{ROH>1Mb}$ of 15.1%, 6.2% and 6.6% for samples of the *Bos taurus* breeds Brown Swiss, Pinzgauer, and Tyrol Grey, respectively. Also, the effective population size estimated for this Nellore sample was approximately 362 animals (**File S1**), which is consistent with the low genome average LD reported by other studies (McKay et al., 2007; Espigolan et al., 2013; Pérez-O'Brien et al., 2014) and indicative of a non inbred population.

3.3. Distribution of chromosome-wise autozygosity

The averages of the chromosome-wise $F_{ROH>0.5Mb}$ across samples are shown in **Figure 3**. Chromosome X exhibited a substantially higher average autozygosity when compared to the autosomes. Importantly, we found no

evidence for a smaller effective population size for the X chromosome in comparison to the autosomal genome (**File S1**). This may be due to the mode of inheritance of the X chromosome, which is hemizygous in the male lineage and therefore more susceptible to bottlenecks and drift even under assumptions of balanced numbers of males and females (Gottipati et al., 2011).

An alternative explanation is that the gene content and the sex-specific copy number of the X chromosome is under stronger selective pressure in comparison to autosomal DNA (Hammer et al., 2010; Deng et al., 2014). In both hypotheses, this higher autozygosity may reflect historical and demographical events. In the early 20th century, when more frequent importation of Nellore cattle to Brazil was initiated, the indigenous herds mainly consisted of descendants from taurine (*Bos taurus*) cattle imported since the late 15th century after the discovery of America (Ajmone-Marsan et al., 2010). In spite of the use of taurine dams for breeding during the early establishment of Nellore cattle in Brazil, the decades that followed were marked by intense backcrossing to Nellore bulls, causing most of the taurine contribution to be swept out from the Nellore autosomal genome (Utsunomiya et al., 2014). However, it is well established that taurine mitochondrial DNA is prevalent in Nellore cattle, as it is a strict maternal contribution (Meirelles et al., 1999). Therefore, the X chromosome may have experienced a greater drift than the autosomal genome due to limited number of founder females, but the levels of taurine introgression still segregating in the X chromosome in this herd remains unclear.

3.4. Identification of common *ROH*

Table 1 presents the results obtained from the *ROH* clustering analysis. The algorithm was robust in respect to gap size between *SNPs*, but substantial differences were observed when the number of consecutive *SNPs* and the number of heterozygous genotypes were modified. Few common *ROH* were identified even when the minimum number of samples in the cluster was 10%, indicating that *ROH* distribution is not uniform across the genome. In fact, despite of the occurrence of 99.98% of the *SNPs* within a *ROH* of at least one individual, only 19.37% markers were encompassed by *ROH* observed in 10%

or more of the samples. This finding is similar to that reported by Ferenčaković et al. (2013b), and is consistent with the stochastic nature of meiotic recombination. This suggests that the ongoing selection for weight, carcass and reproductive traits in this population has not yet created detectable *ROH*-based selection signatures related to production.

The calculations of locus autozygosity were consistent with the cluster analysis using 150 *SNPs* and 2 heterozygous genotypes, regardless of permitted gap size (**Figure 4**). Seven distinct genomic regions, four of them on chromosome X, presented strong hotspots of autozygosity, where over half of the samples ($n = 639$) contained a significant *ROH*. The common *ROH* on the X chromosome are difficult to be discussed as they span several millions of bases, encompassing hundreds of genes and making functional explorations unfeasible. Besides, the assembly status of X chromosome is poorer than the autosomal ones. Hence, we focused on the three autosomal regions on chromosomes 4, 7 and 12. The three regions were relatively short, ranging from 0.73 to 1.43 Mb. For this range of *ROH* length, the expected number of generations since the common ancestor is estimated between 35 and 69 (Howrigan et al., 2011). Assuming a cattle generation interval of 5 years, these inbreeding events may have occurred between 175 and 345 years ago. Although this estimate does not account for birth date and overlapping generations, these remote autozygosity events are likely to predate the foundation of the Nellore breeding programs, and therefore expected to be related to natural selection, random drift or population bottlenecks.

The most autozygous locus was found at chromosome 7:51605639-53035752. This region was previously reported in genome-wide scans for signatures of selection in cattle through the comparison of *Bos taurus* and *indicus* breeds via F_{ST} analysis (Bovine HapMap Consortium, 2009; Porto-Neto et al., 2013) and was detected as a *ROH* hotspot in an analysis of three taurine and indicine breeds each (Sölkner et al., 2014). This region has been implicated in the control of parasitemia in cattle infected by *Trypanosoma congolense* (Hanotte et al., 2003), and is orthologous to the human chromosome segment 5q31-q33, known as the Th2 cytokine gene cluster, which has been shown to be implicated in the control of allergy and resilience against infectious diseases such as malaria (Rihet et al., 1998; Garcia et al., 1998; Hernandez-Valladares

et al., 2004; Flori et al., 2013) and leishmaniasis (Jeronimo et al., 2007). The region also flanks *SPOCK1*, a candidate gene for puberty both in humans (Liu et al., 2009) and cattle (Fortes et al., 2010). Although fertility and resistance to infectious diseases are candidate biological drivers of this *ROH* hotspot, the gene and the phenotype underlying this putative selection signature are unknown.

The common *ROH* at 12:28433881-29743057 identified in the present study also overlaps a common *ROH* hotspot (Sölkner et al., 2014) and a region of divergent selection between *Bos taurus* and *Bos indicus* cattle (Porto-Neto et al., 2013; Gautier et al., 2009), and the segment encompasses the human orthologue *BRCA2*, involved in Fanconi anemia in humans (Howlet et al., 2002). A signature of selection nearby the 4:46384250-47113352 region detected here has also been reported by Gautier & Naves (2011), but the genes involved and the selective pressure remain uncharacterized.

4. Conclusions

We used high-density *SNP* genotypes to successfully characterize autozygosity in Nellore cows under artificial selection for reproductive, carcass and weight traits. We have shown that, although the massive use of relatively few sires and artificial insemination has generated long stretches of homozygous haplotypes in the genomes of over 70% of these animals, inbreeding levels were considerably low in this population. We also found few genomic regions with high homozygosity across individuals, suggesting that the ongoing selection for reproductive, weight and carcass traits in this population is not very intensive or too recent to have left selection signatures in the form of *ROH*. Furthermore, the current common breeding practices of avoiding inbreeding in the mating schemes are antagonistic to additive trait selection, making it hard to maintain *ROH* signatures in the herds. The three candidate regions under selection identified herein were likely to be contributions from remote ancestors, predating the foundation of the Nellore breeding programs. The selective pressure effects and the genes involved in these regions should be subject of future investigation.

5. Acknowledgement

We thank Guilherme Penteado Coelho Filho and Daniel Biluca for technical assistance in sample acquisition. We also thank to Fernando Sebastian Baldi Rey for the manuscript revision and pertinent suggestions. This research was supported by: National Counsel of Technological and Scientific Development (CNPq - <http://www.cnpq.br/>) (process 560922/2010-8 and 483590/2010-0); and São Paulo Research Foundation (FAPESP - <http://www.fapesp.br/>) (process 2013/15869-2 and 2014/01095-8). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Mention of trade name proprietary product or specified equipment in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the authors or their respective institutions.

6. References

- Ajmone-Marsan, P., Garcia, J. F., Lenstra, J. A., and The Globaldiv Consortium. (2010). On the origin of cattle: how aurochs became cattle and colonized the world. *Evol. Anthropol.* 19(4), 148-157.
- Bjelland, D. W., Weigel, K. A., Vukasinovic, N., Nkrumah, J. D. (2013). Evaluation of inbreeding depression in Holstein cattle using whole-genome *SNP* markers and alternative measures of genomic inbreeding. *J. Dairy Sci.* 96(7), 4697-4706.
- Bovine HapMap Consortium. (2009). Genome-wide survey of *SNP* variation uncovers the genetic structure of cattle breeds. *Science.* 324(5926), 528-532.

Curik, I., Ferenčaković, M., Sölkner, J. (2014). Inbreeding and runs of homozygosity: A possible solution to an old problem. *Livestock Science*. 166, 26-34.

Deng, X., Berletch, J. B., Nguyen, D. K., Disteché, C. M. (2014). X chromosome regulation: diverse patterns in development, tissues and disease. *Nat. Rev. Genet.* 15(6), 367-378.

Espigolan, R., Baldi F., Boligon, A. A., Souza, F. R., Gordo, D. G., Tonussi, R. L., Cardoso, D. F., Oliveira, H. N., Tonhati, H., Sargolzaei, M., Schenkel, F. S., Carneiro, R., Ferro, J. A., Albuquerque, L. G. (2013). Study of whole genome linkage disequilibrium in Nelore cattle. *BMC Genomics*. 14,305.

Falconer, D. S., Mackay, T. F. C. (1996). Introduction to Quantitative Genetics. *Addison-Wesley Longman*. 4 ed. (Harlow, Essex, UK), 464.

Ferenčaković, M., Hamzić, E., Gredler, B., Solberg, T.R., Klemetsdal, G., Curik, I., Sölkner, J. (2013a). Estimates of autozygosity derived from runs of homozygosity: empirical evidence from selected cattle populations. *Journal of Animal Breeding and Genetics*. 130, 286-293.

Ferenčaković, M., Sölkner, J., Curik, I. (2013b). Estimating autozygosity from high-throughput information: effects of SNP density and genotyping errors. *Genet. Sel. Evol.* 45,42.

Flori, L., Kumulungui, B., Aucan, C., Esnault, C., Traore, A. S., Fumoux, F., Rihet, P. (2003). Linkage and association between Plasmodium falciparum blood infection levels and chromosome 5q31-q33. *Genes Immun.* 4, 265-268.

Fortes, M. R. S., Reverter, A., Zhang, Y., Collis, E., Nagaraj, S. H., Jonsson, N. N., Prayaga, K. C., Barris, W., Hawken, R. J. (2010). Association weight matrix for the genetic dissection of puberty in beef cattle. *Proc. Natl. Acad. Sci.* 107(31), 13642-136427.

Garcia, A., Marquet, S., Bucheton, B., Hillaire, D., Cot, M., Fievet, N., Dessein, A. J., Abel, L. (1998). Linkage analysis of blood Plasmodium

falciparum levels: interest of the 5q31-q33 chromosome region. *Am. J. Trop. Med. Hyg.* 58(6), 705-709.

Gautier, M., Flori, L., Riebler, A., Jaffrézic, F., Laloé, D., Gut, I., Moazami-Goudarzi, K., Foulley, J. L. (2009) A whole genome Bayesian scan for adaptive genetic divergence in West African cattle. *BMC Genomics*.10, 550.

Gautier, M., Naves, M. (2011). Footprints of selection in the ancestral admixture of a New World Creole cattle breed. *Mol. Ecol.* 20(15), 3128-3143.

Gibson, J., Morton, N. E., Collins, A. (2006). Extended tracts of homozygosity in outbred human populations. *Hum. Mol. Genet.* 15(5), 789-795.

Gottipati, S., Arbiza, L., Siepel, A., Clark, A. G., Keinan, A. (2011). Analyses of X-linked and autosomal genetic variation in population-scale whole genome sequencing. *Nat. Genet.* 43(8), 741-743.

Hammer, M. F., Woerner, A. E., Mendez, F. L., Watkins, J. C., Cox, M. P., Wall, J. D. (2010). The ratio of human X chromosome to autosome diversity is positively correlated with genetic distance from genes. *Nat. Genet.* 42(10), 830-831.

Hanotte, O., Ronin, Y., Agaba, M., Nilsson, P., Gelhaus, A., Horstmann, R., Sugimoto, Y., Kemp, S., Gibson, J., Korol, A., Soller, M., Teale, A. (2003). Mapping of quantitative trait loci controlling trypanotolerance in a cross of tolerant West African N'Dama and susceptible East African Boran cattle. *Proc. Natl. Acad. Sci.* 100(13), 7443-7448.

Hernandez-Valladares, M., Rihet, P., ole-MoiYoi, O. K., Iraqi, F. A. (2004). Mapping of a new quantitative trait locus for resistance to malaria in mice by a comparative mapping approach with human chromosome 5q31-q33. *Immunogenetics.* 56, 115-117.

Hill, W.G. (1975). Linkage disequilibrium among multiple neutral alleles produced by mutation in finite population. *Theor. Popul. Biol.* 8, 117-126.

Howlett, N. G., Taniguchi, T., Olson, S., Cox, B., Waisfisz, Q., De Die-Smulders, C., Persky, N., Grompe, M., Joenje, H., Pals, G., Ikeda, H., Fox, E. A., D'Andrea, A. D. (2002). Biallelic inactivation of *BRCA2* in Fanconi anemia. *Science*. 297(5581), 606-609.

Howrigan, D. P., Simonson, M. A., Keller, M. C. (2011). Detecting autozygosity through runs of homozygosity: a comparison of three autozygosity detection algorithms. *BMC Genomics*. 12, 460.

Jeronimo, S. M., Holst, A. K., Jamieson, S. E., Francis, R., Martins, D. R., Bezerra, F. L., Ettinger, N. A., Nascimento, E. T., Monteiro, G. R., Lacerda, H. G., Miller, E. N., Cordell, H. J., Duggal, P., Beaty, T. H., Blackwell, J. M., Wilson, M. E. (2007). Genes at human chromosome 5q31.1 regulate delayed-type hypersensitivity responses associated with *Leishmania chagasi* infection. *Genes Immun*. 8(7), 539-51.

Keller, M. C., Visscher, P. M., Goddard, M. E. (2011). Quantification of inbreeding due to distant ancestors and its detection using dense single nucleotide polymorphism data. *Genetics*. 189(1), 237-49.

Kim, E. S., Cole, J. B., Huson, H., Wiggans, G. R., Van Tassell, C. P., Crooker, B. A., Liu, G., Da, Y., Sonstegard, T. S. (2013). Effect of artificial selection on runs of homozygosity in U.S. Holstein cattle. *PLoS One*. 8(11), e80813.

Lencz, T., Lambert, C., DeRosse, P., Burdick, K. E., Morgan, V., Kane, J. M., Kucherlapati, R., Malhotra, A. K. (2007). Runs of homozygosity reveal highly penetrant recessive loci in schizophrenia. *Proc. Natl. Acad. Sci.* 104(50), 19942-19947.

Liu, Y. Z., Guo, Y. F., Wang, L., Tan, L. J., Liu, X. G., Pei, Y. F., Yan, H., Xiong, D. H., Deng, F. Y., Yu, N., Zhang, Y. P., Zhang, L., Lei, S. F., Chen, X. D., Liu, H. B., Zhu, X. Z., Levy, S., Papasian, C. J., Drees, B. M., Hamilton, J. J., Recker, R. R., Deng, H. W. (2009). Genome-wide association analyses identify SPOCK as a key novel gene underlying age at menarche. *PLoS Genet*. 5(3), e1000420.

MacArthur, D. G., Balasubramanian, S., Frankish, A., Huang, N., Morris, J., Walter, K., Jostins, L., Habegger, L., Pickrell, J. K., Montgomery, S. B., Albers, C. A., Zhang, Z. D., Conrad, D. F., Lunter, G., Zheng, H., Ayub, Q., DePristo, M. A., Banks, E., Hu, M., Handsaker, R. E., Rosenfeld, J. A., Fromer, M., Jin, M., Mu, X. J., Khurana, E., Ye, K., Kay, M., Saunders, G. I., Suner, M. M., Hunt, T., Barnes, I. H., Amid, C., Carvalho-Silva, D. R., Bignell, A. H., Snow, C., Yngvadottir, B., Bumpstead, S., Cooper, D. N., Xue, Y., Romero, I. G., 1000 Genomes Project Consortium, Wang, J., Li, Y., Gibbs, R. A., McCarroll, S. A., Dermitzakis, E. T., Pritchard, J. K., Barrett, J. C., Harrow, J., Hurles, M. E., Gerstein, M. B., Tyler-Smith, C. (2012). A systematic survey of loss-of-function variants in human protein-coding genes. *Science*. 335, 823.

McKay, S. D., Schnabel, R. D., Murdoch, B. M., Matukumalli, L. K., Aerts, J., Coppieters, W., Crews, D., Neto, E. D., Gill, C. A., Gao, C., Mannen, H., Stothard, P., Wang, Z., Van Tassell, C. P., Williams, J. L., Taylor, J. F., Moore, S. S. (2007). Whole genome linkage disequilibrium maps in cattle. *BMC Genet.* 8, 74.

McQuillan, R., Leutenegger, A. L., Abdel-Rahman, R., Franklin, C. S., Pericic, M., Barac-Lauc, L., Smolej-Narancic, N., Janicijevic, B., Polasek, O., Tenesa, A., Macleod, A. K., Farrington, S. M., Rudan, P., Hayward, C., Vitart, V., Rudan, I., Wild, S. H., Dunlop, M. G., Wright, A. F., Campbell, H., Wilson, J. F. (2008). Runs of homozygosity in European populations. *Am. J. Hum. Genet.* 83(3), 359-372.

Meirelles, F. V., Rosa, A. J. M., Lobo, R. B., Garcia, J. M., Smith, L. C., Duarte, F. A. M. (1999). Is the American zebu really *Bos indicus*? *Genet. Mol. Biol.* 22(4), 543-546.

Pérez O'Brien, A. M., Mészáros, G., Utsunomiya, Y. T., Sonstegard, T. S., Garcia, F. J., Van Tassell, C. P., Carvalheiro, R., Da Silva, M. V. B., Sölkner, J. (2014). Linkage disequilibrium levels in *Bos indicus* and *Bos taurus* cattle using medium and high density *SNP* chip data and different minor allele frequency distributions. *Livest. Sci.* 166, 121-132.

- Porto-Neto, L. R., Sonstegard, T. S., Liu, G. E., Bickhart, D. M., Da Silva, M. V., Machado, M. A., Utsunomiya, Y. T., Garcia, J. F., Gondro, C., Van Tassell, C. P. (2013). Genomic divergence of zebu and taurine cattle identified through high-density *SNP* genotyping. *BMC Genomics*. 14, 876.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller, J., Sklar, P., De Bakker, P. I., Daly, M. J., Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81(3), 559-575.
- Purfield, D. C., Berry, D. P., McParland, S., Bradley, D. G. (2012). Runs of homozygosity and population history in cattle. *BMC Genet.* 13, 70.
- Rihet, P., Traoré, Y., Abel, L., Aucan, C., Traoré-Leroux, T., and Fumoux, F. (1998). Malaria in Humans: *Plasmodium falciparum* Blood Infection Levels Are Linked to Chromosome 5q31-q33. *Am. J. Hum. Genet.* 63, 498-505.
- Sölkner, J., Ferenčaković, M., Karimi, Z., Perez O'Brien, A.M., Mészáros, G., Eaglen, S., Boison, S.A., and Curik, I. (2014). Extremely Non-uniform: Patterns of Runs of Homozygosity in Bovine Populations. *Proceedings, 10th World Congress of Genetics Applied to Livestock Production*.
- Sved, J. A. (1971). Linkage disequilibrium and homozygosity of chromosome segments in finite populations. *Theor. Popul. Biol.* 2, 125-141.
- Szpiech, Z. A., Xu, J., Pemberton, T. J., Peng, W., Zollner, S., Rosenberg, N. A., and Li, J. Z. (2013). Long runs of homozygosity are enriched for deleterious variation. *The American Journal of Human Genetics*. 93, 90-102.
- Tenesa, A.; Navarro, P.; Hayes, B. J.; Duffy, D. L.; Clarke, G. M.; Goddard, M. E.; Visscher, P. M. (2007). Recent human effective population size estimated from linkage disequilibrium. *Genome Res.* 17(4), 520-526.
- Uimari, P. and Tapio, M. (2011). Extent of linkage disequilibrium and effective population size in Finnish Landrace and Finnish Yorkshire pig breeds. *J. Anim. Sci.* 89, 609-614.
- Utsunomiya, Y. T., Bomba, L., Lucente, G., Colli, L., Negrini, R., Lenstra, J. A., Erhardt, G., Garcia, J. F., Ajmone-Marsan, P., and European Cattle

Genetic Diversity Consortium. (2014). Revisiting AFLP fingerprinting for an unbiased assessment of genetic structure and differentiation of taurine and zebu cattle. *BMC Genet.* 15(1), 47.

VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91, 4414-4423.

VanRaden, P. M., Olson, K. M., Wiggans, G. R., Cole, J. B., Tooker, M. E. (2011). Genomic inbreeding and relationships among Holsteins, Jerseys, and Brown Swiss. *J. Dairy Sci.* 94(11), 5673-5682.

Weir, B. S.; Hill, W.G. (1980). Effect of mating structure on variation in linkage disequilibrium. *Genetics.* 95, 477-488.

Xue, Y., Chen, Y., Ayub, Q., Huang, N., Ball, E. V., Mort, M., Phillips, A. D., Shaw, K., Stenson, P. D., Cooper, D. N., Tyler-Smith, C., 1000 Genomes Project Consortium. (2012). Deleterious- and disease-allele prevalence in healthy individuals: insights from current predictions, mutation databases, and population-scale resequencing. *Am. J. Hum. Genet.* 91(6), 1022-1032.

CAPÍTULO 3 – CONSIDERAÇÕES FINAIS

Os programas de melhoramento genético animal são métodos importantes para o desenvolvimento da pecuária moderna. Assim, distintos métodos foram criados para efetuar avaliações fenotípicas e genotípicas com o objetivo de selecionar os rebanhos e os animais geneticamente superiores com produção padronizada. Essa padronização na produção dos animais pode ser alcançada por meio de diversos eventos: i. Acasalamento dirigido entre animais aparentados, isto é, por meio da seleção artificial, ii. Seleção natural, iii. Deriva genética ou iv. Gargalo populacional. Esses eventos contribuem com o aumento da autozigosidade na população, isto é, o aumento do número de animais com alelos idênticos por descendência (*IBD*). Via de regra, a ocorrência desses eventos leva ao aumento da endogamia através do surgimento de regiões autozigotas nos genomas dos indivíduos. As implicações de tais circunstâncias foram descritas no Capítulo 1 dessa dissertação.

A ferramenta mais robusta da atualidade para a identificação dos fragmentos do genoma animal idênticos por descendência é a estimativa do coeficiente de endogamia baseado nas corridas de homoziguidade (*ROH*), que é denominado (F_{ROH}). No Capítulo 2 dessa dissertação, dados genômicos foram empregados para calcular os níveis de autozigosidade baseada em corridas de homoziguidade (*ROH*) a partir de 1.278 fêmeas da raça Nelore genotipadas com mais de 777 mil *SNPs*, para identificar regiões autozigotas possivelmente associadas à seleção natural, domesticação, fertilidade, evolução e adaptação dos bovinos.

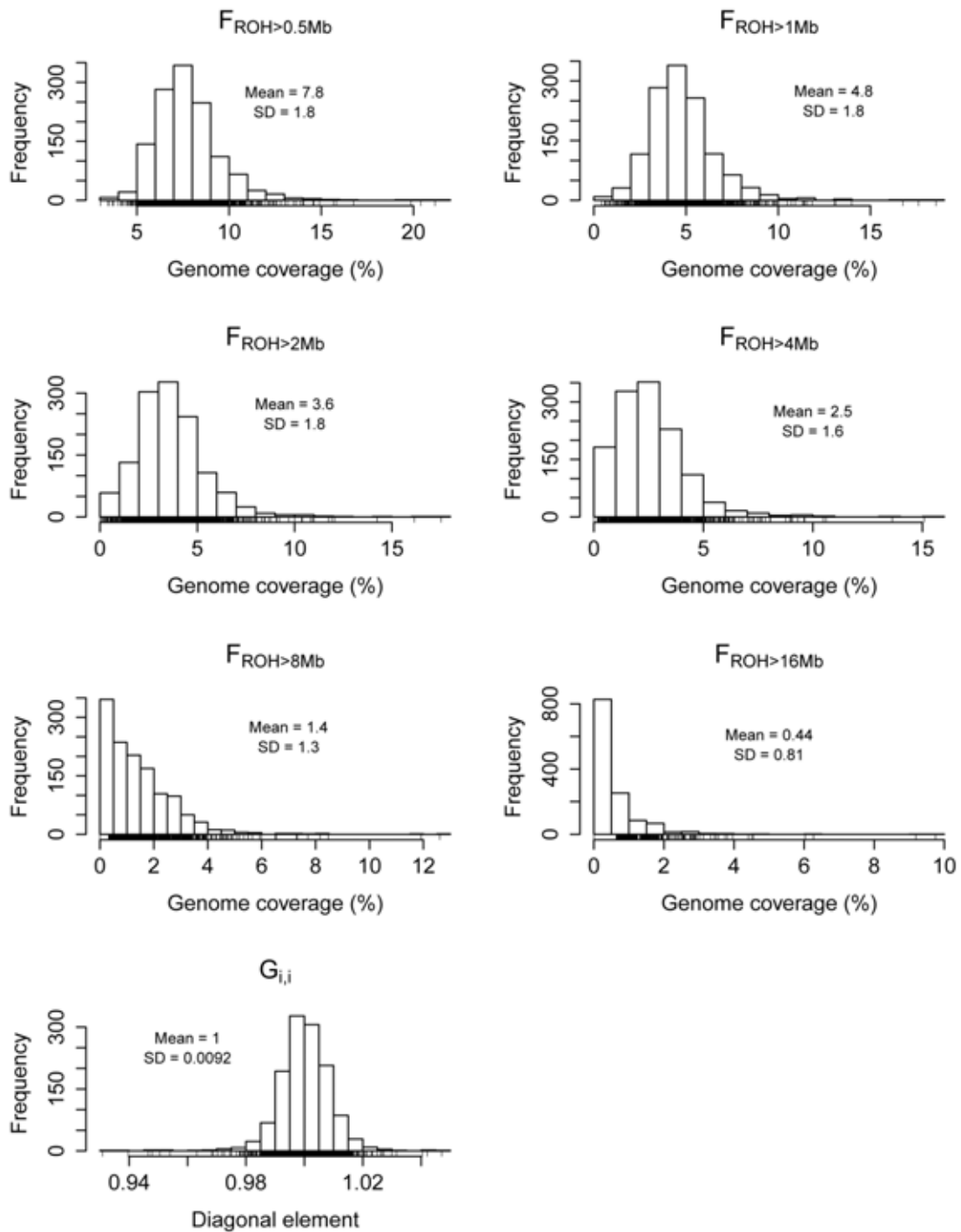
O baixo nível de endogamia genômica do rebanho estudado, apesar do uso de poucos touros fundadores e do uso massivo da inseminação artificial, pode ter sido responsável pela geração de longos trechos de haplótipos homozigotos nos genomas de mais de 70% destes animais. Os resultados deste trabalho apontaram para o aumento da endogamia global na população, devido à permanência de fragmentos longos de *ROH* em 70% dos indivíduos

analisados, indicando que a seleção artificial realizada através do programa de melhoramento genético ao longo de aproximadamente 20 anos (com foco em características reprodutivas, de peso e de carcaça) não foi intensa o suficiente, ou então foi muito recente, para ter deixado assinaturas de seleção na forma de fragmentos de *ROH* comuns.

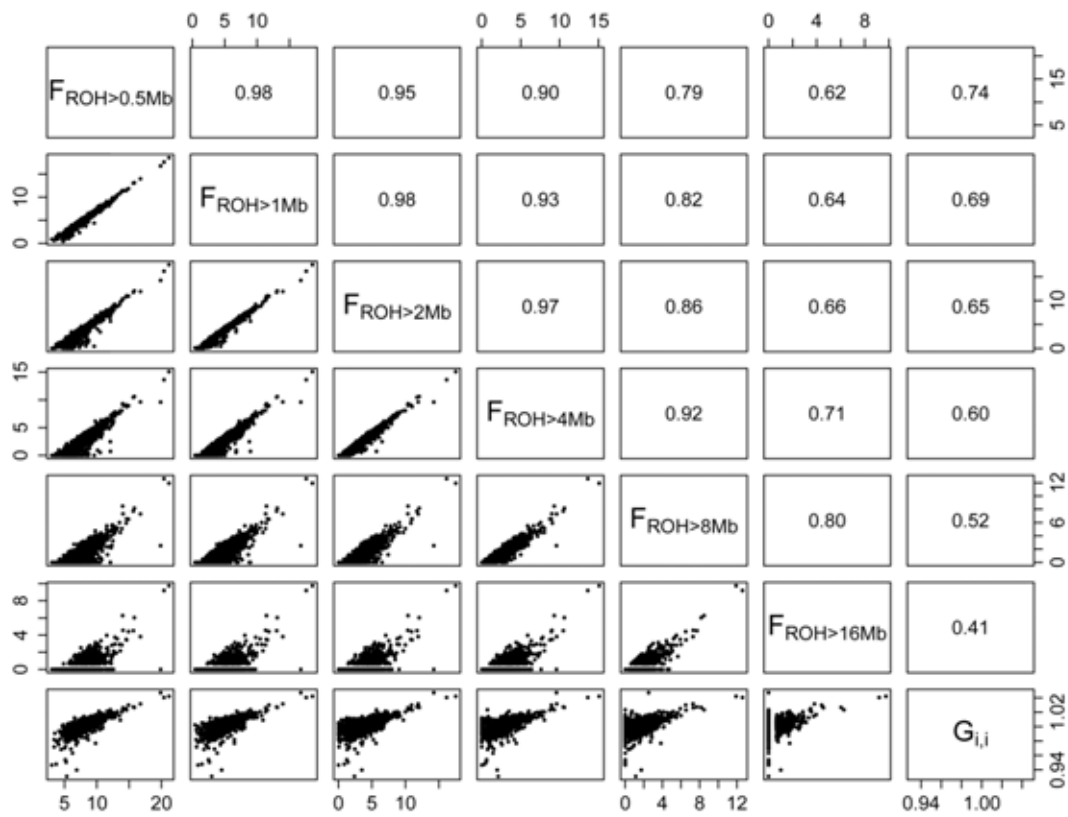
Entretanto, as três regiões de *ROH* comuns ou ilhas de homozigosidade, encontradas nos cromossomos 4, 7 e 12, podem ser referentes as contribuições de antepassados remotos que estejam relacionadas com o processo de seleção natural e a adaptação desses animais, ou seja, à fatos ocorridos anteriormente ao início do programa de melhoramento genético da raça Nelore.

Dessa forma, o presente estudo abriu perspectivas para a realização de análises mais detalhadas, que permitam a mensuração da influência da autozigosidade genômica causada pela pressão seletiva sobre alelos específicos em bovinos da raça Nelore.

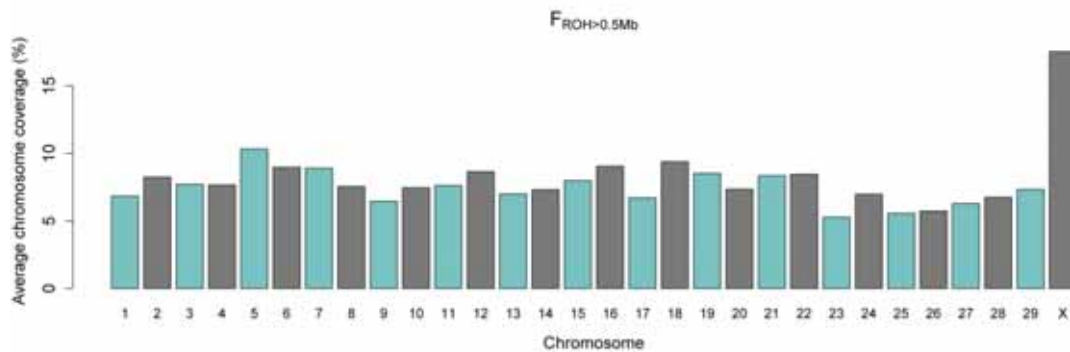
APÊNDICES



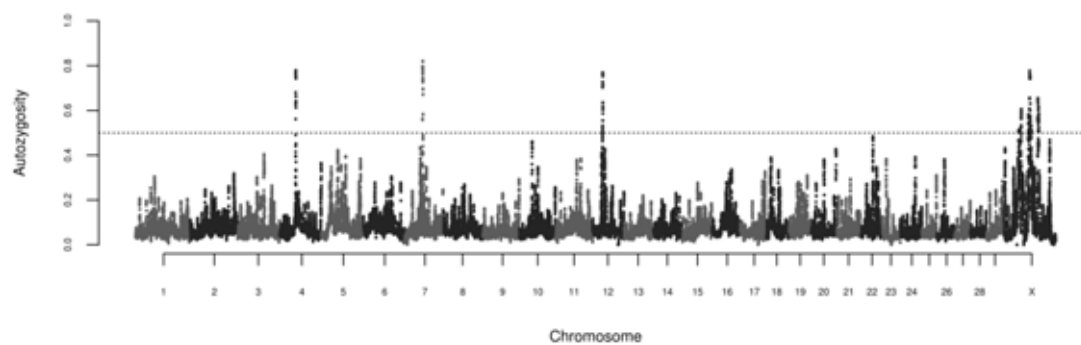
APÊNDICE A – Figure 1. Frequency distribution of percentage of autosomal genome coverage by runs of homozygosity (F_{ROH}) of different minimum lengths (>0.5 , >1 , >2 , >4 , >8 and >16 Mb) and diagonal elements of the realized genomic relationship matrix ($G_{i,i}$).



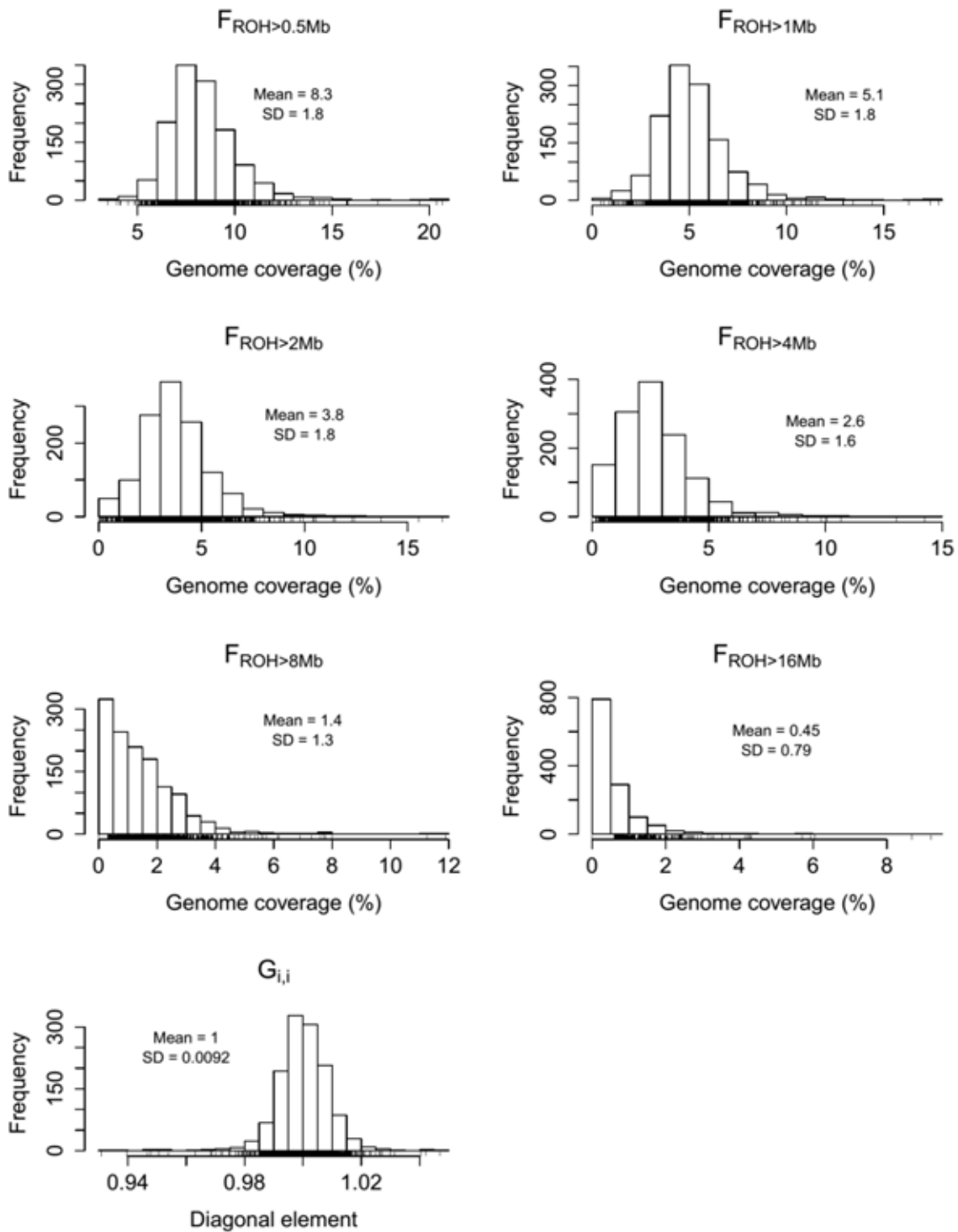
APÊNDICE B – Figure 2. Scatterplots (lower panel) and correlations (upper panel) of percentage of autosomal genome coverage by runs of homozygosity (F_{ROH}) of different minimum lengths (>0.5, >1, >2, >4, >8 and >16 Mb) and diagonal elements of the realized genomic relationship matrix ($G_{i,i}$).



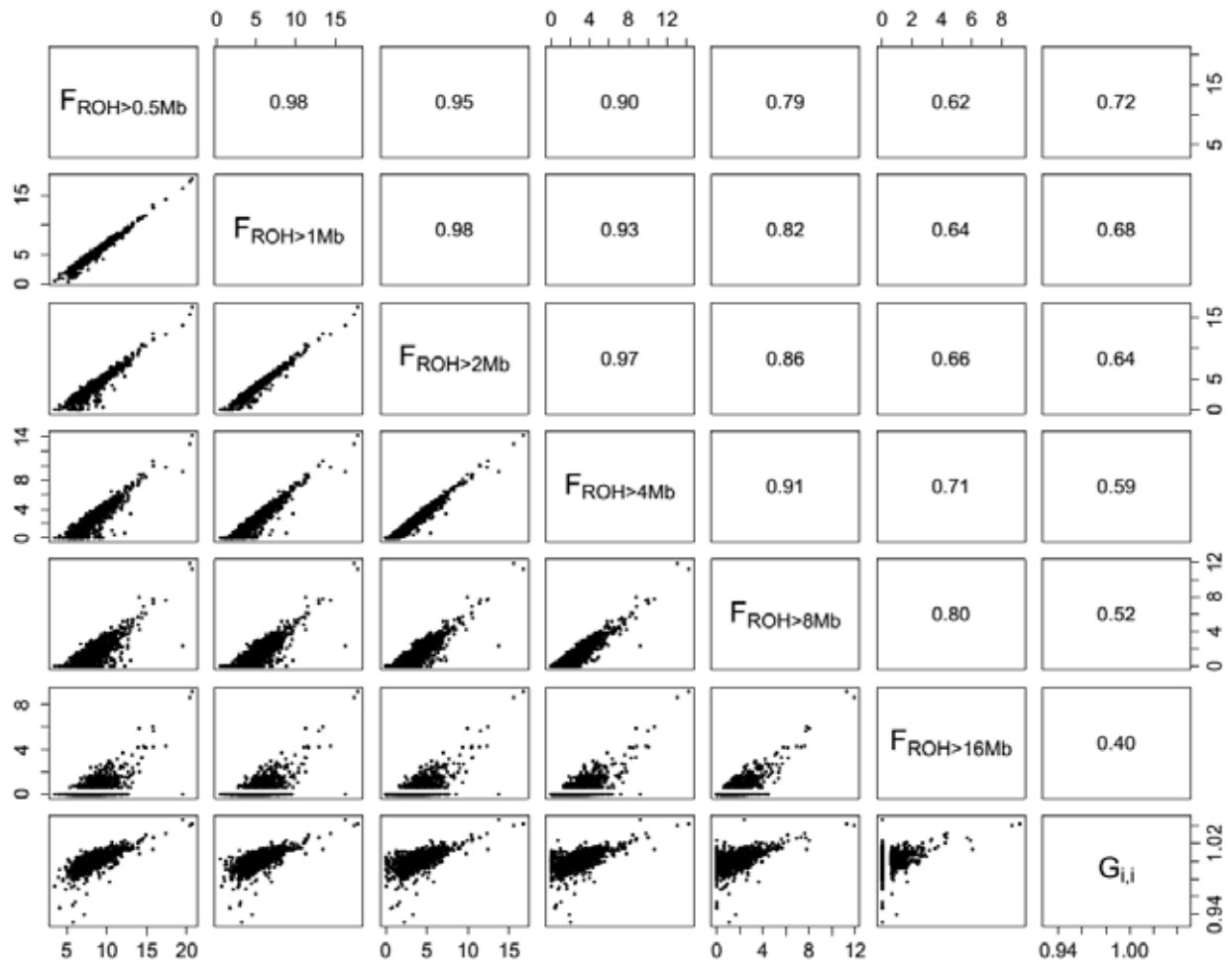
APÊNDICE C – Figure 3. Barplot of average percentage of chromosome coverage by runs of homozygosity (F_{ROH}) of minimum length of 0.5 Mb.



APÊNDICE D – Figure 4. Manhattan plot of genome-wide locus autozygosity in Nellore cows. The dashed line represents the 50% threshold.



APÊNDICE E – Figure S1. Frequency distribution of percentage of the total genome (autosomes + X) coverage by runs of homozygosity (F_{ROH}) of different minimum lengths (>0.5, >1, >2, >4, >8 and >16 Mb) and diagonal elements of the realized genomic relationship matrix ($G_{i,i}$).



APÊNDICE F – Figure S2. Scatterplots (lower panel) and correlations (upper panel) of percentage of total genome (autosomes + X) coverage by runs of homozygosity (F_{ROH}) of different minimum lengths (>0.5 , >1 , >2 , >4 , >8 and >16 Mb) and diagonal elements of the realized genomic relationship matrix ($G_{i,i}$).

APÊNDICE G – Table 1. Detection of common runs of homozygosity according to different number of consecutive *SNPs*, percentage of animals, gap size and number of heterozygous genotypes.

Gap size	30 <i>SNPs</i>				150 <i>SNPs</i>				Heterozygotes
	10%	20%	25%	50%	10%	20%	25%	50%	
100 Kb	437	106	57	9	186	29	12	1	0
	471	288	183	29	365	91	47	7	2
500 Kb	479	126	76	13	193	32	14	1	0
	768	334	214	45	375	96	50	7	2

APÉNDICE H – Table S1. Parameter estimates for the nonlinear regression models by chromosome

Chromosome	Number of pairwise comparisons	$\hat{\alpha}$	$SE(\hat{\alpha})$	$\hat{\beta}$	$SE(\hat{\beta})$	$\hat{N}_e$
1	7,154,155	4.052	0.003	1318.138	1.493	329,534
2	5,723,254	3.715	0.003	1479.120	1.779	369,780
3	6,155,446	4.321	0.004	1078.419	1.456	269,604
4	4,783,055	4.021	0.004	1597.914	2.206	399,478
5	5,128,268	4.085	0.004	1092.324	1.488	273,081
6	6,575,478	3.935	0.003	1233.294	1.362	308,323
7	5,172,412	3.594	0.003	1500.191	1.873	375,047
8	6,322,519	3.965	0.003	1251.309	1.440	312,827
9	5,380,061	3.710	0.003	1625.961	1.959	406,490
10	4,091,772	3.790	0.004	1840.789	2.673	460,197
11	4,564,182	3.768	0.003	1590.636	2.151	397,659
12	3,738,278	4.183	0.004	1382.465	2.224	345,616
13	3,506,642	3.852	0.004	1648.760	2.566	412,190
14	4,679,140	4.079	0.004	1326.500	1.753	331,625
15	3,428,608	4.048	0.004	1612.331	2.515	403,082
16	3,788,210	3.872	0.004	1553.210	2.245	388,302
17	4,031,419	4.901	0.005	798.770	1.468	199,692
18	2,969,957	3.578	0.004	1957.124	3.248	489,281
19	2,251,363	4.303	0.006	1599.284	3.219	399,821
20	3,133,031	3.969	0.004	1499.503	2.449	374,875
21	3,259,818	3.820	0.004	1395.949	2.355	348,987
22	2,583,455	4.138	0.005	1656.291	2.995	414,072
23	2,399,662	4.586	0.006	1718.008	3.305	429,502
24	2,822,059	3.850	0.005	1558.057	2.878	389,514
25	2,038,428	5.382	0.010	811.048	2.404	202,762
26	2,356,368	3.914	0.005	2029.060	3.904	507,265
27	1,832,751	4.534	0.007	1797.444	3.986	449,361
28	1,915,010	4.932	0.007	1279.928	3.002	319,982
29	1,830,764	4.319	0.007	1658.245	3.729	414,561
5 replicates of						
All autosomes	29,000,000	4.099	0.004	1448.766	1.940	361,786
X	7,341,954	3.884	0.003	554.181	0.781	277,090

APÊNDICE I – Table S2. Parameter estimates for the nonlinear regression model for 5 replicates of 1 million randomly sampled r^2 values from each autosomal chromosome.

Autosome replicate	$\hat{\alpha}$	$\hat{\beta}$	$\hat{N}_e$
1	4.097	1447.460	361.865
2	4.101	1446.598	361.649
3	4.108	1444.103	361.026
4	4.096	1448.812	362.203
5	4.097	1448.766	362.192
Average	4.100	1447.148	361.787
Stdev	0.005	1.941	0.485

APÊNDICE J – Methods for effective population size estimation.

Effective population size (N_e) was estimated based on its approximate relationship with linkage disequilibrium (LD). More precisely, the expected squared correlation of allele frequencies at a pair of loci (r^2) can be related to N_e as follows:

$$E(r^2) = \frac{1}{(\alpha + kN_e c)} + var(r)$$

Where c is the linkage distance (in Morgans) between the loci, α and k are constants, and $var(r)$ is the chance disequilibrium variance introduced by experimental sampling. The analysis of autosomal markers ($k = 4$) under the assumption of absence of mutation ($\alpha = 1$) and sampling error ($var(r) = 0$) leads to the formula discovered by Sved (1971). Likewise, taking sampling error ($var(r) = 1/n$, where n is the number of haplotypes), mutation ($\alpha = 2$) and sex chromosomes into account ($k = 4$ for autosomes and $k = 2$ for sex chromosomes) leads to other known extensions of the problem (Hill, 1975; Weir & Hill, 1980). Based on this formula, Tenesa et al. (2007) introduced a nonlinear regression model to estimate N_e from genome-wide *SNP* data:

$$y_i = \frac{1}{(\alpha + \beta c_i)} + e_i$$

Where $y_i = r^2 - 1/n$ is the LD for marker pair i corrected for the number of haplotypes, β is a parameter representing kN_e , c_i is the genetic distance between markers (in Morgans), and e_i is an error term. First, due to computational complexity, we used *PLINK* v1.07 (Purcell et al., 2007) to calculate pairwise r^2 only for *SNPs* presenting $MAF > 0.05$ that were no more than 3 Mb apart. This distance constraint considered the rapid LD decay reported for the Nellore genome (McKay et al., 2007; Espigolan et al., 2013; Pérez-O'Brien et al., 2014). Following Uimari & Tapio (2010), only LD values of $0.01 < r^2 < 0.99$ were included, as the point estimates of N_e are infinite at $r^2 = 0$ and $r^2 = 1$. Second, the approximation $c_i = 10^{-8}d_i$ was used, where d_i is the inter-marker distance in base pairs for *SNP* pair i , under the assumption of $1\text{Mb} \approx 1\text{cM}$. Third, the model above was fitted for each chromosome and α and β were estimated using nonlinear least squares in *R* v3.1.1 (<http://www.r-project.org/>). Chromosomal N_e was derived as $\hat{N}_e = \hat{\beta}/4$ for autosomes and $\hat{N}_e = \hat{\beta}/2$ for the X chromosome. Finally, as the total number of computed r^2 values was too large to be regressed using available hardware (approximately 120 million), the genome-wide $\hat{N}_e$ was derived by fitting the nonlinear model using 1 million randomly sampled r^2 values from each autosome (summing up 29 million observations). Five replicates were processed in order to measure sampling variance.