

UNIVERSIDADE ESTADUAL PAULISTA “JULIO MESQUITA FILHO”

FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS

CÂMPUS DE JABOTICABAL

**ESTUDO DO DESEQUILÍBRIO DE LIGAÇÃO E ESTIMATIVA DO
TAMANHO EFETIVO EM UMA POPULAÇÃO DA RAÇA GIR
SELECIONADA PARA CRESCIMENTO PÓS-DESMAMA.**

Alejandra Maria Toro Ospina

Zootecnista

2017

D
I
S
S

O
S
P
I
N
A

A.
M.
T.

2
0
1
7

UNIVERSIDADE ESTADUAL PAULISTA “JULIO MESQUITA FILHO”

FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS

CÂMPUS DE JABOTICABAL

**ESTUDO DO DESEQUILÍBRIO DE LIGAÇÃO E ESTIMATIVA DO
TAMANHO EFETIVO EM UMA POPULAÇÃO DA RAÇA GIR
SELECIONADA PARA CRESCIMENTO PÓS-DESMAMA.**

Alejandra Maria Toro Ospina

Orientador: Prof. Dr. Josineudson Augusto II de Vasconcelos Silva

**Dissertação apresentada à
Faculdade de Ciências Agrárias e
Veterinárias – Unesp, Câmpus de
Jaboticabal, como parte das
exigências para a obtenção do
título de Mestre em Genética e
Melhoramento Animal**

2017

T686e Toro Ospina, Alejandra Maria
Estudo do desequilíbrio de ligação e estimativa do tamanho efetivo em uma população da raça gir selecionada para crescimento pós-desmama / Alejandra Maria Toro Ospina. -- Jaboticabal, 2017
iii, 41 p. : il. ; 29 cm

Dissertação (mestrado) - Universidade Estadual Paulista, Faculdade de Ciências Agrárias e Veterinárias, 2017
Orientador: Josineudson Augusto Il de Vasconcelos Silva
Banca examinadora: Lenira el Faro Zadra, Rusbel Raul Aspilcueta Borquis
Bibliografia

1. Seleção genômica. 2. Haplótipos. 3. Deriva genética. 4. Endogamia. I. Título. II. Jaboticabal-Faculdade de Ciências Agrárias e Veterinárias.

CDU 636.2:636.082

Ficha catalográfica elaborada pela Seção Técnica de Aquisição e Tratamento da Informação – Serviço Técnico de Biblioteca e Documentação - UNESP, Câmpus de Jaboticabal.

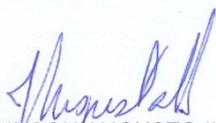
CERTIFICADO DE APROVAÇÃO

TÍTULO: ESTUDO DO DESEQUILÍBRIO DE LIGAÇÃO E ESTIMATIVA DO TAMANHO EFETIVO EM UMA POPULAÇÃO DA RAÇA GIR SELECIONADA PARA CRESCIMENTO PÓS-DESMAMA

AUTOR: ALEJANDRA MARIA TORO OSPINA

ORIENTADOR: JOSINEUDSON AUGUSTO II DE VASCONCELOS SILVA

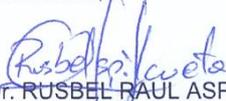
Aprovado como parte das exigências para obtenção do Título de Mestre em GENÉTICA E MELHORAMENTO ANIMAL, pela Comissão Examinadora:



Prof. Dr. JOSINEUDSON AUGUSTO II DE VASCONCELOS SILVA
Departamento de Nutrição e Melhoramento Animal / FMVZ / UNESP - Botucatu



Pesquisadora Dra. LENIRA EL FARO ZADRA
IZ / Sertãozinho, SP



Prof. Dr. RUSBEL PAUL ASPILCUETA BORQUIS
Universidade Federal da Grande Dourados / Dourados, MS

Jaboticabal, 24 de fevereiro de 2017.

DADOS CURRICULARES DO AUTOR

Alejandra Maria Toro Ospina – Filha de Whiston Líder Toro Garcia e Maria del Carmen Ospina Sanchez, nascida no dia 11 de janeiro de 1992, no município de Trujillo Valle del Cauca, Colombia. Zootecnista, graduada pela Faculdade universidade de Santa Rosa de Cabal, Colombia, em dezembro de 2014, com trabalho de grado titulado “ efecto del reemplazo del 50% de plasma seminal em toros Gyr durante la congelacion sobre el acrosoma y la endosmosis del espermatozoide” sob a orientação do Dr. Prof. German Dario Saavedra. Em julho de 2014 realizou um estagio na área de genética, na Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista (UNESP) sob orientação do Dr. Prof. Humberto Tonhati. Em agosto de 2015, ingressou no programa de Genética e Melhoramento Animal na Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista (UNESP), Câmpus de Jaboticabal, sob orientação do Dr. Prof. Josineudson Augusto II de Vasconcelos Silva. Durante o curso recebeu apoio financeiro da CAPES.

DEDICATORIA

Dedico este trabalho a meus pais Whiston Lider e Maria del Carmen, pelo amor e apoio incondicional em todas minhas metas, que graças ao esforço deles hoje posso cumprir, são meus anjos que sempre me protegem.

A meu irmão Whiston Edwin por ser meu segundo pai e por acreditar sempre em mim, muito obrigada pelos ensinamentos, apoio, amor em minha vida toda.

A minhas irmãs Luz Enith e Adriana Isabel por sempre estar para mim, ser meu exemplo a seguir, para incentivar-me a seguir os meus sonhos, e saber que tudo com amor da família é possível.

A meus Sobrinhos Jhon Stiven, Natalia e Ana Lucia, pelo amor que sempre me dão, são a motivação para cada dia ser melhor e ajuda-los no futuro.

OS AMO MUITO, SÃO MINHA FORÇA PARA IR EM FRENTE COM NOVAS METAS...

AGRADECIMENTOS

Agradeço ao Professor Augusto pela oportunidade de fazer mestrado com ele, pelos ensinamentos, paciência e apoio em todo. Muito obrigada por acreditar em mim.

Á minha amiga e companheira de trabalho Amanda pela ajuda, paciência e amizade em todo o processo de aprendizagem.

Á todas as pessoas que de alguma maneira ajudaram no desenvolvimento do trabalho, colegas do professor agosto.

Ao doutor Danisio Munari e Fernando Baldi pelas correções e considerações para melhorar meu trabalho.

A universidade UNESP, á CAPES pela concessão da bolsa de estudos.

Á minha amiga Valentina por sua amizade e companhia em Brasil por estar sempre para me ajudar em todo.

Ana Cristina, por ser minha professora e amiga, muito obrigada pelos ensinamentos, apoio, paciência e amizade.

Ao Professor Guilherme Venturini pela ajuda e contribuições a meu trabalho.

A meus companheiros de sala Eduardo, camilla, pablo, Ricardo pelo companheirismo no tempo de mestrado.

A meus companheiros do programa de GMA, Lucio, Guilherme, André, Diercles pelas contribuições a meu trabalho.

SUMÁRIO

CAPITULO 1 – CONSIDERAÇÕES GERAIS	1
1. Introdução	1
2. Revisão de literatura	2
2.1 Desequilíbrio de ligação	2
2.2 Tamanho efetivo da população	5
2.3 Painel de alta densidade	7
2.4 Imputação de Genótipos	7
3. Referências bibliográficas	8
CAPÍTULO 2 - DESEQUILIBRIO DE LIGAÇÃO E TAMANHO EFETIVO EM POPULAÇÃO DA RAÇA GIR SELECIONADA PARA CRESCIMENTO PÓS-DESMAMA	14
Introdução.....	14
Objetivo	15
Material e métodos.....	15
Resultados e Discussão.....	19
Conclusões.....	36
Referências bibliográficas	36

DESEQUILÍBRIO DE LIGAÇÃO E TAMANHO EFETIVO EM POPULAÇÃO DA RAÇA GIR SELECIONADA PARA CRESCIMENTO PÓS-DESMAMA

RESUMO - O objetivo deste estudo foi estimar o desequilíbrio de ligação (r^2) nas distâncias de 25-50kb, 50-100kb, 100-500kb, 0,5-1Mb e o tamanho efetivo (N_e) nas gerações 0, 5, 10, 15, 20 em população da raça Gir selecionada para crescimento pós-desmama. Os animais utilizados no presente estudo foram provenientes do rebanho fechado do Instituto de Zootecnia, Sertãozinho, SP. Foram obtidos os genótipos de 155 animais com o painel BovineDL 33kb e 18 com painel HD imputado onde realizou-se controle de qualidade (CQ) para alelo de menor frequência (MAF) $< 0,02$ e *call rate* $< 0,1$. Depois do CQ permaneceram 27.236 SNPs e 155 animais do painel de 33 kb e 732.962 SNPs e 173 animais do painel HD Imputado. As análises de r^2 foram realizadas pelo programa Plink e programa estatístico R Studio e o N_e por meio do DL. Os resultados das distâncias 25-50kb, 50-100kb, 100-500kb e 0,5-1Mb do r^2 para o painel 33kb foram iguais a 0,29, 0,25, 0,16 e 0,032 respectivamente, e 0,35, 0,29, 0,18, 0,032 para o painel HD imputado demonstrando que o DL permaneceu nas distâncias menores a 100kb, decaindo com o aumento das distâncias. Estes resultados foram maiores aos descritos na literatura para animais zebuínos, sugerindo como causa os segmentos longos de haplótipos que compartilham os animais aparentados. O N_e foi igual a 9, 17, 24, 30 e 30 animais nas gerações 0, 5, 10, 15, 20, observa-se que o N_e é maior na geração 20, com 30 animais, e decai drasticamente a partir da 5 geração com 17 animais, e sendo de 9 animais a última geração, um tamanho pequeno para uma população. Os valores encontrados neste estudo mostram alto DL e baixo N_e , provavelmente pelo sistema de seleção e a estrutura da população da raça Gir avaliada, que apresenta alto nível de endogamia, perda da variabilidade genética, uso intensivo de pequeno número de reprodutores, conduzindo a diminuição da deriva genética da população, ocasionando dificuldades na seleção dos animais.

Palavras-chave: Seleção genômica, haplótipos, deriva genética, endogamia.

LINKAGE DISEQUILIBRIUM AND EFFECTIVE SIZE ON POPULATION OF GIR ZEBU BREED SELECTED FOR POST-WEANING WEIGHTS.

SUMMARY - The aim of this study was to estimate the linkage disequilibrium (r^2) at distances of 25-50kb, 50-100kb, 100-500kb, 0,5-1Mb and the effective population size (N_e) in generations 0, 5, 10, 15, 20 in population of the selected Gir for yearling growth. The animals used in this study were from the closed herd Animal Science Institute, Sertãozinho, SP. the genotypes of 155 animals were obtained with BovineDL 33kb and 18 animals of panel HD, where quality control was held (QC) for minor allele frequency (MAF) <0.02 and call rate <0.1 . After QC remained 27,236 SNPs and 155 animals to panel 33 kb, 732.962 SNPs and 173 the panel HD imputation. The r^2 analyzes were performed by Plink program and R Studio statistical program and N_e through the DL. The results of r^2 for distances 25-50kb, 50-100kb, 100-500kb and 0,5-1Mb were equal to 0.29, 0.25, 0.16 and 0.032, respectively, showing that the DL remained in smaller distances 100kb, decreasing with increasing distances. These results were higher than those reported in the literature for Zebu animals, suggesting a cause to long haplotype segments that share the related animals. N_e is equal to 9, 17, 24, 30 and 30 in the generations 0, 5, 10, 15, 20, it is observed that N_e is higher in generation 20 with 30 animals and decays sharply from 5 Generation 17 animals, and with 9 animals the latest generation, small size for a population. The values found in this study to DL and N_e , explain the selection system and the structure of the population of Gir evaluated, which has a high level of inbreeding, loss of genetic variability, intensive small number of players, leading to decreased drift population genetics, causing difficulties in the selection of the next generations.

Keywords: genomic selection, haplotypes, genetic drift, inbreeding.

CAPITULO 1 – CONSIDERAÇÕES GERAIS

1. Introdução

Bos primigenius indicus, conhecido como Zebu, é o bovino mais popular entre os países tropicais, algumas das raças mais representativas desta espécie são: Nelore, Guzerá, Gir, Indubrasil (ASSOGIR, 2015). O Gir provavelmente seja a raça zebuína mais antiga do planeta, segundo a literatura sagrada hinduísta. Morfologicamente, sua antiguidade também se manifesta pela conformação craniana, é a única raça bovina com chifres voltados para baixo e para trás, e de crânio ultra convexo, no mundo. O Gir possui grande popularidade na Índia e em países tropicais, principalmente pela notável mansidão e aptidão leiteira e tem sido objeto de muitos estudos (ABCGIL, 2015).

A raça Gir chegou ao Brasil em 1911, porém tornou-se figura comum nas propriedades agropecuária, em meados da década de 1960, para atender o mercado leiteiro, diversos produtores iniciaram projetos de selecionar as fêmeas Gir de aptidão leiteira com finalidade de encontrar os animais mais produtivos. Desta forma, dividiu-se a raça em duas linhagens: linhagem de leite e linhagem de corte.

O rebanho Gir do Instituto de Zootecnia começou em 1976, e com base na progênie resultante nascida em 1981, foi realizada a primeira seleção para crescimento pos-desmame (RAZOOK et al., 1998). A única introdução de material genético de outras linhagens do Gir entrou em 1986, para aumentar a variabilidade, mas depois se torne um rebanho fechado por mais de 40 anos.

Estudos realizados na raça Gir, selecionada para características de crescimento, descreveram tendências positivas dos efeitos genéticos materno para ganho de peso pré-desmame e direto pós-desmame obtendo resposta genética significativa na seleção direta dos animais Gir (EUCLIDES FILHO, 2000; KNACKFUSS et al, 2006).

A aplicação dos marcadores SNPs no melhoramento animal, área conhecida como seleção genômica, encontra-se condicionada primeiramente ao estudo de desequilíbrio de ligação destes marcadores. O desequilíbrio de ligação (DL) é definido como as associações não aleatórias entre alelos em loci diferentes dentro de determinada população, apresentando grande importância no contexto

de mapeamento fino do QTL (locos de caracteres quantitativos) e na estimação de parâmetros populacionais (TENESA et al., 2007).

Outra utilização importante do DL consiste na estimação do tamanho efetivo da população (N_e) em estudo. O DL ocorre exclusivamente da deriva genética para loci neutro desvinculado em população isolada com acasalamento ao acaso (HILL, 1981). Este fenômeno pode ser utilizado para estimar o tamanho efetivo (N_e) da população explorando a variação do DL ou a correlação das frequências entre locis de genes (r).

O sequenciamento do genoma do Gir no Brasil levou quatro anos e foi concluído em 2015, envolvendo pesquisadores da Embrapa, da Universidade Federal de Minas Gerais e da Fiocruz-Minas. O feito tem importância histórica, já que foi o primeiro sequenciamento do genoma de um mamífero feito por equipe 100% brasileira (EMBRAPA, 2015).

A imputação de genótipos tem sido explorada como estratégia na redução dos custos da implementação da seleção genômica (HUANG et al., 2012; KHATKAR et al., 2012). O objetivo principal da imputação é genotipar indivíduos referências com painel de alta densidade de marcadores (HDP), e candidatos à seleção com painéis de baixa densidade (LDP), os quais têm custo menor, e, então, predizer os genótipos de locos faltantes com informações provenientes da população referência para uniformizar o genoma (HOWIE et al., 2009; HAYES et al., 2011; LI et al., 2011).

Com intuito de contribuir com estudos na área de seleção genética e genômica da raça Gir, o objetivo do presente projeto foi analisar o desequilíbrio de ligação e tamanho efetivo, utilizando painéis de alta densidade de SNPs em população de bovinos da raça Gir selecionada para corte.

2. Revisão de literatura

2.1 Desequilíbrio de ligação

O termo desequilíbrio de ligação (DL) foi utilizado pela primeira vez em 1960 por Lewontin e Kojima e emerge da análise das frequências de pares de alelos durante a formação de gametas, por comparação entre as frequências observadas

com aquelas esperadas sob independência, obtidas com base nas frequências de cada um dos alelos que compõem esses pares. Quando as frequências gaméticas observadas em determinada população são idênticas às esperadas obtidas pelo produto das frequências alélicas envolvidas, diz-se que os locus estão em equilíbrio de ligação (LE). Quando não ocorre, têm-se a falta de equilíbrio de ligação, ou seja, o desequilíbrio de ligação (DL). Assim, quando há desvio nas frequências observadas dos haplótipos quando comparadas às esperadas tem-se o desequilíbrio de ligação (LEWONTIN & KOJIMA, 1960; LEWONTIN, 1964).

Outros autores definiram DL como sendo a associação não aleatória de alelos de diferentes locus, que ocorre quando dois genes estão suficientemente próximos no genoma, sendo difícil a recombinação durante a meiose, e os segmentos do cromossomo são conservados de uma geração para outra (ARDLIE et al., 2002; BOHMANOVA et al., 2010). Do ponto de vista estatístico, o conceito de DL pode ser definido como medida estatística da associação de alelos de diferentes locus.

O interesse dos pesquisadores no estudo do DL nas espécies cresceu rapidamente na década de 1980 (HEDRICK, 1987), devido a utilidade na genética de populações, por permitir visão detalhada da evolução e divergências genéticas das espécies, além de servir como base para o mapeamento de genes (SLATKIN, 2008).

O desequilíbrio de ligação é parâmetro importante na biologia evolutiva e genética, por conter informações sobre eventos ocorridos na história de uma população, ou seja, variações ocorridas com o passar das gerações e que afetam o DL, como a seleção, mutação e recombinação; também informações sobre haplótipos transmitidos entre as gerações (DEKKERS, 2004) e alterações no DNA, como polimorfismos e mutações. O DL em determinada região genômica reflete a história da seleção natural, a conversão do gene, mutação e outros fatores. O DL diminui com o passar das gerações em todo o genoma, com o aumento da distância entre os marcadores e de diferentes magnitudes da distância entre marcador e QTL (ROOS et al., 2008).

Os bovinos *Bos primigenius indicus* adaptaram-se em ambientes rústicos e foram selecionados para produção em áreas tropicais e subtropicais, enquanto animais *Bos primigenius taurus* em regiões temperadas. As diferenças no processo

histórico de formação das subespécies bovinas desencadearam diferenças em regiões do genoma (PÉREZ O'BRIEN et al., 2014). A quantidade de DL que existe nas regiões genômicas é utilizada como ferramenta para rastrear pressões seletivas e identificar diminuição do DL entre populações bovinas (PEREZ O'BRIEN et al., 2014). Além disso, o DL está relacionado com o tamanho efetivo das populações, e pode fornecer conhecimentos sobre a história e a diversidade genética. E, com as estimativas da fase de ligação se pode determinar a existência da relação entre as raças ou em uma mesma população e utilizá-las como populações de referências em programas de seleção genômica (BIEGELMEYER et al., 2016).

O DL tem sido amplamente estudado no genoma de raças taurinas. Zhu et al. (2013) realizaram estudo com 640 touros da raça Simental para estimar DL em cromossomos autossômicos com diferentes distâncias entre SNPs aleatórios, e encontraram valores médios de r^2 iguais a 0,30, 0,16 e 0,08 quando a separação entre os SNPs foi de 0 a 25 kb, de 50 a 100 kb e de 0,5 a 1 Mb, respectivamente. Biegelmeyer et al. (2016) obtiveram os genótipos de 391 animais da raça Hereford com painel de 41.241 SNPs e 2.044 animais da raça Braford com painel de 41.207 SNP e estimaram r^2 superior a 0,2 e 0,3, respectivamente, para 34 e 25% de marcadores adjacentes em Hereford, e 26 e 17% em Braford.

Por outro lado, Beghain et al. (2013) avaliaram 30 touros da raça Blonde d'Aquitaine com painel de 36.923 SNPs, e a média obtida do r^2 foi 0,205 (\pm 0,262). Os autores também verificaram DL alto ($r^2 > 0,8$) e DL com r^2 superior a 0,2 dentro das regiões genômicas de até 720 e 724 kb, respectivamente, revelando que o nível de DL diminui rapidamente com a distância física entre os SNPs.

Segundo Lu et al. (2012), os valores de r^2 são similares entre diferentes raças, afirmação com base em estudo de DL nas raças Angus e Charolês, considerando painel de 31.073 e 32.088 SNPs, respectivamente. As estimativas de r^2 obtidas foram 0,29, 0,23 e 0,19 na raça Angus e 0,22, 0,16 e 0,12 na raça Charolês, quando os intervalos de distância entre os marcadores foram de 0 a 30 kb, 30 a 70 kb e 70 a 100 kb.

Em raças leiteiras também ocorre semelhança nos resultados de r^2 , Roos et al. (2008) estudando touros das raças Holstein-Friesian australiana, Angus, Nova Zelândia Friesian e Jersey obtiveram média de r^2 igual a 0,35, 0,25, 0,22, 0,14 em distâncias entre marcadores de 10, 20, 40, 100 e 1.000 kb, respectivamente,

verificando rápido decréscimo no DL com o aumento das distâncias nos animais *Bos primigenius taurus*.

Em zebuínos, estudos de DL foram realizados nas raças Nelore e Gir. O estudo com a raça Nelore mostrou diminuição do r^2 com o aumento da distância física entre os marcadores de 0,34 (1 kb) a 0,11 (100 kb). Os cromossomos BTA1, BTA27, BTA28 e BTA29 apresentaram menores níveis de DL a qualquer distância entre os marcadores (ESPIGOLAN et al. 2013).

Na raça Gir, a média de r^2 em todos os autossomos foi de 0,21 e 0,68, demonstrando tendência de diminuição exponencial com o aumento da distância física. As médias de r^2 obtidas na raça Gir foram 0,20, 0,18, 0,14 e 0,11 para intervalos de 0 a 0,1 Mb, 0,2 a 0,0 Mb, 0,0 a 0,5 Mb e 0,0 a 1,0 Mb, respectivamente (SILVA et al. 2010). As raças zebuínas têm proporção superior de alelos de menor frequência (MAF) e nível de DL menor em relação às raças taurinas (SILVA et al. 2010).

2.2 Tamanho efetivo da população

O tamanho efetivo da população (N_e) é definido como a dimensão de uma população idealizada (FISHER, 1930; WRIGHT, 1931). Eventos como mutação, seleção, migração e recombinação interferem no valor de N_e , com o qual é possível identificar a quantidade e distribuição da variação genética presente em determinada população. Além disso, o N_e ajuda a prever a perda e a distribuição da variação genética, as probabilidades de fixação de alelos benéficos ou prejudicial e aptidão e sobrevivência de uma população (WANG, 2005).

O estudo do DL sobre diferentes distâncias nos permite estimar N_e em vários momentos. Segundo relatos da literatura, o N_e em animais *Bos primigenius taurus* foi superior a 50.000 antes da domesticação, e em torno de 1.000 a 2.000 após a domesticação, diminuindo para cerca de 100 nos últimos anos em muitas raças (GODDARD, 1992).

Conforme o valor de N_e , espera-se padrão onde existe algum DL nas populações recentes de animais taurinos, considerando distâncias inferiores a 10 cM, porém o DL aumenta lentamente à medida que diminui a distância e, eventualmente, atinge alto r^2 em distâncias muito pequenas (GODDARD & HAYES, 2012).

Os bovinos *Bos primigenius indicus* têm maior heterozigosidade e, portanto, no passado deve ter apresentado maior N_e comparado aos *Bos primigenius taurus* (The BovineHapMap Consortium, 2009).

Segundo a FAO (2016), uma amostra de 25 machos e 50 fêmeas é recomendada como sendo o mínimo para um programa de conservação de animais vivo. Cálculo obtido de forma a resultar em perda de menos de 1% da possível variação genética presente na população original (SMITH, 1983). Com relação ao valor mínimo de N_e de uma população, existe discussão com relação ao preconizado por Frankel & Soule (1981) que descrevem tamanho efetivo mínimo absoluto de uma população ser 50 animais, tamanho considerado necessário para a sobrevivência de populações do jardim zoológico, de espécies selvagens, onde as estratégias de reprodução podem ser controladas de perto.

O decréscimo drástico do valor de N_e pode representar perda de variabilidade e impedir a formação das próximas gerações. Entretanto, o N_e pode diferir entre populações e raças (BIEGELMEYER et al., 2016).

Lee et al. (2011) observaram divergência na tendência do N_e entre duas populações bovinas. Na raça Bradford encontraram N_e igual a 220, com aumento nas duas últimas gerações, e na raça Hereford, o N_e obtido foi 153, com quadro de declínio acelerado do valor a partir das quatro últimas gerações. Um estudo na raça coreana Hanwoo Korean, o N_e obtido foi de 98, ocorrido nas três últimas gerações (LEE et al., 2011).

Para zebuínos, os valores verificados de N_e são menores em relação aos citados anteriormente. Faria et al. (2009) estimaram N_e nas raças Nelore, Gir e Guzará para o período de 1979 a 1998. Os valores diminuíram de 85 para 68 no Nelore, de 70 para 45 no Gir e manteve-se praticamente constante (104) no Guzará. Outro estudo na raça Gir (Mocha) na região Nordeste do Brasil mostrou decréscimo do N_e na população, no período de 1954 a 2005, sendo extremamente pequeno na última geração, com 20 animais (MUNIZ et al., 2012).

O N_e pode ser calculado com diferentes fontes de informação (informação demográfica, de linhagem ou de dados moleculares) vários indicadores de diversidade genética podem alterar o N_e como a deriva genética através de mudanças temporais nas frequências alélicas, o aumento da homozigose pela endogamia ou a taxa em que os alelos originais são perdidos (LEROY et al., 2013).

2.3 Painel de alta densidade

Em estudos de desequilíbrio de ligação de bovinos, foi pouco utilizado o painel de alta densidade, porém ao passar dos anos e com a implementação da seleção genômica, ficou demonstrado que a utilização dos painéis de SNPs de alta densidade, superior a 700k seria a melhor decisão para estudos de associação ampla do genoma (GWAS) para bovinos de corte (MOKRY et al., 2014).

A associação e a aplicação da seleção genômica com as plataformas de obtenção dos genótipos, torna possível investigar DL com alta resolução, possibilitando detectar sinais de associação eficazes, quando utilizado painel de alta densidade de SNPs, ao contrário dos de baixa densidade, que podem ocultar informações (PORTO-NETO et al., 2014).

Para realizar seleção genômica necessita-se utilizar painéis de alta densidade com marcadores espalhados por todo o genoma, com a finalidade de ter marcadores próximos aos QTL que tenham desequilíbrio de ligação. Os painéis de alta densidade têm garantia de alto desequilíbrio de ligação, e permite a explicação da variância genética aditiva das características de interesse.

2.4 Imputação de Genótipos

Para a imputação de genótipos existem programas baseados em dois métodos de imputação: métodos baseados na população, que usam o desequilíbrio de ligação (DL) populacional; e métodos baseados na informação de família, segregação ou na combinação de informações de família, segregação e DL populacional (LI et al., 2009). Os programas mais usados na inferência de genótipos em produção animal e que apresentam boa acurácia de imputação são Impute2 (HOWIE et al., 2009) e Beagle (BROWNING and BROWNING, 2007, 2013), que são baseados no DL populacional, e FImpute (SARGOLZAEI et al., 2014), baseado na informação de família.

Métodos que usam a informação da população geralmente são probabilísticos ou baseado em modelos e exploram o DL entre SNPs próximos considerando as frequências de haplótipos. Neste método, os indivíduos não são considerados parentes, entretanto, o parentesco próximo ainda é capturado entre os indivíduos por meio dos longos haplótipos compartilhados entre eles (BROWNING and

BROWNING, 2009). Para este caso, a acurácia da imputação depende principalmente da densidade do painel e tamanho da população referência.

O FImpute considera o método baseado em família, trabalhando com a genealogia e o parentesco entre os indivíduos. A imputação acurada dos genótipos é realizada com uso de informações de haplótipos curtos e longos e informações de parentes próximos e distantes. A rapidez de processamento do FImpute é devido a sua metodologia de análise determinística, o que facilita seu uso em grandes conjuntos de dados, quando o uso de outros métodos são impraticáveis (SARGOLZAEI et al., 2014).

3. Referências bibliográficas

ABCGIL. ASSOCIACAO BRASILEIRA DOS CRIADORES DE GIR LEITERO. Disponível em: <<http://girleiteiro.org.br/>> Acesso em 12 de jan de 2016.

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

ASSOGIR. ASSOCIACAO BRASILEIRA DOS CRIADORES DE GIR. Disponível em: <<http://assogir-brasil.blogspot.com.br/>> Acesso em 13 de jan de 2015.

BEGHAIN, J.; BOITARD, S.; WEISS, B.; BOUSSAHA, M.; GUT, I.; ROCHA, D. Genome-wide linkage disequilibrium in the Blonde d'Aquitaine cattle. *Journal of Animal Breeding and Genetics*, 130, 294-302, 2013.

BIEGELMEYER, P.; GULIAS-GOMES, C.C.; CAETANO, A.R.; STEIBEL, Y.J.P. Linkage disequilibrium, persistence of phase and effective population size estimates in Hereford and Braford cattle. *BMC Genetics*, 17, 1-12, 2016.

BOHMANOVA, J.; SARGOLZAEI, M.; SCHENKEL, F. Research article Characteristics of linkage disequilibrium in North. *BMC Genomics.*, 11:421, 2010.

BROWNING, B.L. and S.R. BROWNING, Efficient multilocus association testing for whole genome association studies using localized haplotype clustering. *Genet Epidemiol.* 31(5): p. 365-75, 2007.

Browning, B.L. and S.R. Browning, A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *Am J Hum Genet.* 84(2): p. 210-23, 2009.

DEKKERS, J.C.M. Commercial application of marker and gene assisted selection. *J. Anim. Sci.*, 82, E-Suppl:E313-328, 2004.

EMBRAPA. Empresa Brasileira de Pesquisa Agropecuária. Disponível em <<https://www.embrapa.br/web/portal/busca-de-noticias/-/noticia/2634883/pesquisadores-brasileiros-sequenciam-o-genoma-do-gir-leiteiro>> Acesso em 17 de set de 2015.

ESPIGOLAN, R.; BADI, F.; BOLIGON, A.A.; et al. Study of whole genome linkage disequilibrium in Nelore cattle. *BMC Genomics*, 14, 305, 2013.

EUCLIDES FILHO, K.; SILVA, L.O.C.; ALVES, R.G.O.; FIGUEIREDO, G.R. Tendência genética na raça Gir. *Pesquisa Agropecuária Brasileira.* 35, 787-791, 2000.

FAO. In situ conservation of livestock and poultry. Disponível em <<http://www.fao.org/docrep/004/t0559e/t0559e04.htm>> Acesso em 17 de set. 2016.

FARIA, F.J.C.; VERCESI FILHO, A.E.; MADALENA, F.E.; JOSAHKIAN, L.A. Pedigree analysis in the Brazilian Zebu breeds. *Journal of Animal Breeding and Genetics.* 126, 148-153, 2009.

FISHER R.A. The genetical theory of natural selection. Oxford University Press, , p.308, 1930.

FRANKEL, O.H.; SOULÉ, M.E. Conservation and evolution. Cambridge University Press, Cambridge. U.K., 327p, 1981.

GODDARD, M.E. Optimal effective population size for the global population of black and white dairy cattle. *J. Dairy Sci.* 75, 2902–2911, 1992.

GODDARD, M.E.; HAYES, B.J. Genome-Wide Association Studies and Linkage Disequilibrium in Cattle. *Bovine Genomics*. Chapter 13, p.194-195, 2012.

HAYES BJ, BOWMAN PJ, DAETWYLER HD, KIJAS JW, VAN DER WERF JHJ: Accuracy of genotype imputation in sheep breeds. *Anim Genet.* 43:72–80, 2011.

HEDRICK, P.W. Genetic disequilibrium measures. *Genetics.* 117, 331-341, 1987.

HILL, W.G. Estimation of effective population size from data on linkage disequilibrium. *Genetical Research.* 38, 209-216, 1981.

HOWIE, B. N., P. DONNELLY, AND J. MARCHINI.. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* 5:e1000529, 2009.

HUANG, Y., J. M. HICKEY, M. A. CLEVELAND, and C. MALTECCA.. Assessment of alternative genotyping strategies to maximize imputation accuracy at minimal cost. *Genet. Sel. Evol.* 44:25. 2012

KHATKAR, M. S., G. MOSER, B. J. HAYES, and H. W. RAADSMA. Strategies and utility of imputed SNP genotypes for genomic analysis in dairy cattle. *BMC Genomics* 13:538. 2012.

KNACKFUSS, F.B.; RAZOOK, A.G.; MERCADANTE, M.E.Z. et al. Seleção para peso pós-desmama em um rebanho Gir. 1. Resposta direta e correlacionada em 21 anos de seleção. *R. Bras. Zootec.* 35, 715-725, 2006.

LEE, S.H.; CHO, Y.M.; LIM, D.; et al. Linkage Disequilibrium and Effective Population Size in Hanwoo Korean Cattle. *Asian-Aust. J. Anim. Sci.* 24, 1660-1665, 2011

LEROY, G., TRISTAN, M.H., VERRIER, E., DANVY, S., CHARVOLIN E AND DANCHIN-BURGE C. Methods to estimate effective population size using pedigree data: Examples in dog, sheep, cattle and horse. *Genetics Selection Evolution.*, 45:1, 2013.

LEWONTIN, R.C. The interaction of selection and linkage. I. General considerations; Heterotic Models. *Genetics*, 49, 49-67, 1964.

LEWONTIN, R.C.; KOJIMA, K. The Evolutionary Dynamics of Complex Polymorphisms. *Evolution.* 14, 458-472, 1960.

LI L, LI Y, BROWNING SR, BROWNING BL, SLATER AJ, KONG X, APONTE JL, MOOSER VE, CHISSOE SL, WHITTAKER JC, NELSON MR, EHM MG: Performance of genotype imputation for rare variants identified in exons and flanking regions of genes. *PLoS One*, 6(9):e24945, 2011.

LU, D.; SARGOLZAEI, M.; KELLY, M.; et al. Linkage disequilibrium in Angus, Charolais, and Crossbred beef cattle. *Frontiers in Genetics.* 3, 152, 2012.

MOKRY, F.B.; BUZANSKAS, M.E.; MUDADU, M.A.; et al. Linkage disequilibrium and haplotype block structure in a composite beef cattle breed. *BMC Genomics.* 15(Suppl 7), S6, 2014.

MUNIZ, L.M.S. A raça Gir Mocha na região Nordeste do Brasil: estrutura genética populacional via análise de pedigree. *Arq. Bras. Med. Vet. Zootec.* 64, 1656-1664, 2012.

PÉREZ O'BRIEN, A.M.; UTSUNOMIYA, Y.T.; MÉSZÁROS, G. et al. Assessing signatures of selection through variation in linkage disequilibrium between taurine and indicine cattle. *Genetics, Selection, Evolution.* 46, 19, 2014.

PORTO-NETO, L.R.; KIJAS, J.W.; REVERTER, A. The extent of linkage disequilibrium in beef cattle breeds using high-density SNP genotypes. *Genet Sel Evol.* 24, 46:22, 2014.

RAZOOK, A,G,; FIGUEIREDO, L,A,; BONILHA NETO, L,M,; et al, Selection for yearling weight in Nelore and Guzará zebu breeds: selection applied and response in 15 years of progeny, In: WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 6, 1998, Proceedings,,,, Armidale: NSW, 1998, p,133-136

ROOS, A.P.W.; HAYES, B.J.; SPELMAN, R.J.; GODDARD, M.E. Linkage Disequilibrium and Persistence of Phase in Holstein–Friesian, Jersey and Angus Cattle. *Genetics.* 179, 1503-1512, 2008.

SARGOLZAEI, M.; CHESNAIS, J.P.; SCHENKEL, F.S. A new approach for efficient genotype imputation using information from relatives. *BMC Genomics*, 15, 478, 2014.

SILVA, C. R.; NEVES, H. H. R.; QUEIROZ, S. A.; et al. Extent of linkage disequilibrium in Brazilian Gyr dairy cattle based on genotypes of AI sires for dense SNP markers. In: 9th World Congress on Genetics Applied to Livestock Production, Leipzig, Alemanha. Proceedings of the 9th World Congress on Genetics Applied to Livestock Production, 2010.

SLATKIN, M. Linkage disequilibrium understanding the evolutionary past and mapping the medical future. *Nature reviews genetics.*, 9: 477-485, 2008.

SMITH, J.K. Quantitative versus qualitative research: An attempt to clarify the issue. *Education Researcher*, 12, 6-13, 1983.

TENESA, A.; NAVARRO, P.; HAYES, B.J.; DUFFY, D.L.; CLARKE, G.M.; GODDARD, M.E.; VISSCHER, P.M. Recent human effective population size estimated from linkage disequilibrium. *Genome Research*, 17, 520–526, 2007.

The BovineHapMap Consortium. Bovine Genome Sequencing and Analysis Consortium. The genome sequence of taurine cattle: a window to ruminant biology and evolution. *Science*. 324, 522-8. 2009.

WANG, J. Estimation of effective population sizes from data on genetic markers. *Philosophical Transactions of the Royal Society B. Biological Sciences*. 29, 1395-1409, 2005.

WRIGHT, S. Evolution in Mendelian populations. *Genetics*. 16, 97-159, 1931.

ZHU, M.; ZHU, B.; WANG, Y.H. Linkage Disequilibrium Estimation of Chinese Beef Simmental Cattle Using High-density SNP Panels. *Asian-Australasian Journal of Animal Science*. 26, 772-779, 2013.

CAPÍTULO 2 - DESEQUILÍBRIO DE LIGAÇÃO E TAMANHO EFETIVO EM POPULAÇÃO DA RAÇA GIR SELECIONADA PARA CRESCIMENTO PÓS-DESMAMA

Introdução

O desequilíbrio de ligação (DL) ocorre quando há desvio nas frequências observadas dos haplótipos quando comparadas às esperadas. Se ocorre o aparecimento de novo alelo, por exemplo via mutação, e tem maior frequência, os outros alelos permanecerão associados, exceto na ocorrência de recombinação (GODDARD & HAYES, 2012). O DL é primordial nas análises genômicas, porque rastreia a história das populações para entender eventos passados e futuros para mapear genes que estão associados com características quantitativas que são transmitidas por herança. A seleção, a deriva genética, recombinação e mutação afetam os níveis de DL, sendo utilizados na análises de associação do genoma (GWAS) (SLATKIN 2008).

O tamanho efetivo é definido como a dimensão de uma população idealizada (FISHER 1930; WRIGHT, 1931). O N_e ajuda a explicar a variação genética da população, prevenindo a perda da deriva genética e a sobrevivência de populações. Em bovinos, tamanho efetivo populacional foi estimada a partir da endogamia nas populações. No entanto, a estimativa do tamanho efetivo da população pode ser afetada por vários eventos desconhecidos (HEDRICK, 2005).

Um das maneiras de calcular o N_e com informação genômica é por meio do DL, se o N_e for grande, o ancestral comum provavelmente deve ser muito antigo, somente pequenos segmentos dos cromossomos terão sobrevivido sem recombinação, por outro lado, se os segmentos do cromossomo são longos, estes serão idênticos por descendência e o N_e vai ser pequeno e com DL alto na população (GODDARD & HAYES, 2012).

Em virtude da recente aplicação de seleção genômica e estudos de associação ampla do genoma para maior compreensão do genoma e de características de importância econômica para o melhoramento genético dos bovinos e da carência de estudos básicos com painéis de arranjos de SNPs na

raça Gir, o objetivo do presente estudo foi estimar o desequilíbrio de ligação e o tamanho efetivo da população de Gir selecionada para corte, utilizando marcadores de alta densidade de SNPs.

Objetivo

O objetivo do presente projeto será estudar o desequilíbrio de ligação e estimar tamanho efetivo de uma população de Gir utilizando marcadores de alta densidade de SNPs.

Material e métodos

Animais

Os animais usados nestes experimento foram provenientes do rebanho do Centro Avançado de Pesquisa Tecnológica do Agronegócio (CAPTA) Bovinos de Corte, unidade de pesquisa do Instituto de Zootecnia (IZ), localizado na cidade de Sertãozinho, SP. A raça Gir participou do programa de seleção para crescimento pós desmama das raças zebuínas do IZ.

O processo de reestruturação do rebanho Gir do IZ, tendo em vista a primeira estação de monta, começou em 1976, e com base na progênie resultante nascida em 1981, foi realizada a primeira seleção (RAZOOK et al., 1998). O primeiro procedimento adotado foi introdução de material genético de outras linhagens do Gir, com o intuito de promover maior variabilidade genética no rebanho, evitando altos níveis de endogamia no decorrer dos anos de seleção.

A escolha dos touros fundadores, nascidos entre 1977 e 1980, baseou-se no critério de seleção vigente na época (peso aos 550 dias de idade, P550). Em 1980, montou-se o rebanho Gir com 80 matrizes e, anualmente foram selecionados três tourinhos (equivalente a 10% da zafra) e de 15 a 20 vacas (equivalente a 50-60% do grupo contemporâneo). A partir deste ano, foi estabelecida seleção de touros com base no máximo diferencial de seleção para o peso aos 378 dias (P378) obtido no final da prova de ganho de peso, e seleção de matrizes com base no P550 obtido em recria a pasto.

No arquivo de pedigree, foram incorporadas registros desde o ano de 1968 até 2013, possibilitando acesso à genealogia completa do rebanho Gir, para a análise conjunta com as informações de SNPs dos animais.

Genótipos

O material biológico (sangue ou pelo) dos animais foi coletado após decisão do IZ de vender o rebanho no ano de 2013. Portanto, neste experimento foram agrupados dados de genótipos do único rebanho da raça Gir selecionado para corte no Brasil.

Foram genotipados 18 animais com o painel BovineHD BeadChip da Illumina, abrangendo 772.962 SNPs do genoma bovino, e 155 animais com o painel Bovine LDv4 (33K) da Illumina, o qual tem cobertura de aproximadamente 33.000 SNPs. Os animais genotipados com o painel HD serviram como referência para imputação de genótipos que foi realizada com uso do programa FImpute (SARGOLZAEI et al., 2014). A imputação dos genótipos no FImpute considera o método baseado em família, trabalhando com informações de haplótipos curtos e longos e informações de parentesco entre os indivíduos. O uso da informação de parentes próximos é necessário para imputação de variantes raras (SARGOLZAEI et al., 2014). Previamente, SNPs dos cromossomos sexuais foram excluídos, de forma que trabalhamos apenas com cromossomos autossômicos no processo de imputação de genótipos. SNPs do painel 33K que não estavam presentes no painel HD também foram excluídos.

O controle de qualidade dos genótipos foi realizado após a imputação. Como parâmetros de controle de qualidade dos SNPs considerou-se $MAF < 0,02$ e $call\ rate < 0,1$, e $call\ rate < 0,90$ para amostra.

Dados imputados

Os genótipos imputados do painel 33K (155 animais) foram consolidados nesta etapa como genótipos “verdadeiros”, permanecendo então com 732.463 SNPs para cada animal. Para verificar a acurácia da imputação selecionou-se cada um dos 18 animais em painel HD e retirou-se os SNPs que não estavam presentes

no painel 33K, com intuito de mascarar os genótipos verdadeiros para serem imputados. Após a adequação, foi realizada a imputação por meio de uma análise com todos os animais, inclusive os 155 animais já imputados. Este processo foi realizado para cada um dos 18 animais.

As análises estatísticas para obter o desempenho métrico da imputação foram realizadas para cada um dos 18 animais, comparando os SNPs imputados com o real do HD. Foram calculadas três medidas de performance métrica com o objetivo de capturar diferentes características da acurácia e qualidade da imputação.

O primeiro foi o cálculo da taxa de concordância (P_o), obtida pela percentagem de genótipos imputados e uma probabilidade de acordo (P_c) que correspondem com o genótipo verdadeiro.

Em que:

P_o = taxa de concordância

P_c = probabilidade de acordo

$$P_o = \frac{\sum_{n=1}^N P_{11_n} + \sum_{n=1}^N P_{22_n} + \sum_{n=1}^N P_{33_n}}{N}$$

$$P_c = \frac{N1 * \sum_{j=1}^3 \sum_{n=1}^N P_{1j_n} + N2 * \sum_{j=1}^3 \sum_{n=1}^N P_{2j_n} + N3 * \sum_{j=1}^3 \sum_{n=1}^N P_{3j_n}}{N^2}$$

Este cálculo da taxa de concordância tem sido utilizado como medida de acurácia da imputação (HUANG et al., 2009; CHARLES et al., 2011).

O segundo, usando os mesmos SNPs mascarados foi calculado o índice de qualidade da imputação (IQS) como descrito por Lin et al. (2010), sendo particularmente útil para avaliar a acurácia da imputação de variantes raras e de baixa frequência.

Em que:

$$IQS = \frac{P_o - P_c}{1 - P_c}$$

O terceiro foi estimada a correlação simples (CS) entre cada genótipo imputado e seu genótipo subjacente verdadeiro, e então a média de todas as correlações dos SNPs polimórficos imputados.

Os métodos para estimar a acurácia foram realizadas pelo programa SAS

Análises Estatísticas

Coeficiente de correlação de desequilíbrio de ligação (r^2)

O LD foi medido usando o coeficiente de correlação de desequilíbrio de ligação (r^2) entre dois SNPs, considerando a equação matemática:

$$r^2 = \frac{D^2}{p^1 p^2 q^1 q^2}$$

em que, D é a diferença entre observada e Frequência esperada de um haplótipo, ou desvio; e p^1 , p^2 , q^1 e q^2 referem-se as frequências alélicas baseadas nas frequências dos haplótipos, sendo as frequências dos alelos A_1 , A_2 , B_1 e B_2 , respectivamente.

O r^2 foi obtido pelo programa PLINK v1.90 (comando -- plink.ld) e considerado no cálculo da estatística r^2 para todos os pares de SNPs em cada cromossomo, por meio do programa R Studio (<http://www.r-project.org/>). Em grandes bases de dados, r^2 igual a 1 indica completo DL; r^2 igual a 0 significa que os SNPs estão de equilíbrio de ligação, ou seja, não há LD.

Endogamia Genômica FROH

Convencionalmente, os níveis de endogamia para animais são estimados analisando dados de pedigree. A recente disponibilidade de arrays de SNP de alta densidade para o genoma bovino forneceu Oportunidade de investigação de níveis de autozygosidade com base em corridas de homozigose (ROH). Uma corrida de homozigose é um segmento contínuo de seqüência de DNA sem heterozigocidade no diploide. O análise pelo programa PLINK.

Estimação do tamanho efetivo

A relação entre o tamanho efetivo da população (N_e), taxa de recombinação e DL (r^2) sem mutação está resumida na seguinte equação:

$$r^2 = 1/[4N_e c + 1]$$

onde c é a distância entre os marcadores no mapa de ligação em morgans (SVED, 1971). O mapa da distância de ligação (c) será inferido com base na relação entre

o tamanho físico de cada cromossomo e comprimento do mapa de ligação correspondente.

O tamanho efetivo populacional será estimado como:

$$N_e = \left(\frac{1}{4c}\right) \left(\frac{1}{r^2 - 1}\right)$$

em que, N_e é o tamanho efetivo populacional, o valor de 4 é o número de gerações, e c unidade de centimorgas assumindo que 1Mb igual a 1cM e o r^2 é desequilíbrio de ligação entre os alelos (HAYES et al., 2003; GODDARD et al., 2006). Para este cálculo, foi usado o programa R studio, considerando os 173 animais estudados.

Parâmetros populacionais com informação do pedigree

Os parâmetros populacionais foram calculados com o programa ENDOG v.4.6 (GUTIÉRREZ & GOYACHE, 2005). O coeficiente de endogamia individual (F) foi estimado de acordo com Falconer e Mackay (1996), em que F é igual ao relacionamento genético aditivo (f) entre seus pais, ou coeficiente de coancestralidade. Por meio do aumento individual na endogamia, o tamanho efetivo da população (N_e) foi estimado com base em cada geração $F_t > F_{t-1}$, caracterizando o efeito remoto e fechado da endogamia. Os parâmetros de probabilidade de origem do gene - número efetivo de fundadores (f_e), número efetivo de ancestrais (f_a) e número efetivo de genomas fundadores (f_g) - foram obtidos com uso do programa ENDOG. O f_e é fornecido pela mensuração das contribuições dos fundadores mais influentes. O f_a representa o número mínimo de ancestrais (fundadores ou não) necessários para explicar a diversidade genética completa da população. O f_g é definido como o número de reprodutores que contribuem de forma igual na estrutura populacional e produzem idêntica diversidade genética, sem ocorrer perdas de alelos.

Resultados e Discussão

Informações de SNPs para o painel de 33kb de cada cromossomo autossômico são dadas na Tabela 1. O comprimento total dos cromossomos autossômicos da população de Gir foi 2.535.82 Mb. O maior cromossomo *Bos indicus* autossômica foi o BTA2 (comprimento igual a 158,26 Mb), e o mais curto foi BTA25

(comprimento igual a 42.72 Mb). Neste estudo foram utilizados 27.236 SNPs cobrindo todo o genoma.

O DL foi significativamente diferente entre os cromossomos. Obtendo a média de r^2 para os SNPs separados por intervalos de 25 a 50 kb, 50 a 100 kb, 100 a 500 kb, 0,5 a 1 Mb em cada cromossomo autossômico (Tabela 2). O valor médio de r^2 para a distância de 25 a 50kb nos diferentes cromossomos foi alto para BTA5 e BTA20, sendo 0,405 e 0,435, respectivamente, e baixos para BTA18 (0,153). A média do r^2 na distância 50 a 100kb foi maior para BTA5, BTA20 (0,336 e 0,410) e baixo para o BTA26 com 0,198.

Tabela 1. Informação estatística das análises de SNP com painel de 33k em população da raça Gir selecionada para crescimento pós desmama

Cromossomo	Mb	Número de SNPs	r^2
BTA1	158,26	1554	0,026
BTA2	163,67	1309	0,025
BTA3	121,20	1263	0,026
BTA4	120,08	1185	0,027
BTA5	120,88	1448	0,033
BTA6	119,34	1289	0,029
BTA7	112,52	1141	0,029
BTA8	113,30	1101	0,032
BTA9	105,58	1099	0,033
BTA10	104,04	1044	0,028
BTA11	107,23	1123	0,028
BTA12	91,01	927	0,036
BTA13	84,07	944	0,031
BTA14	84,59	990	0,032
BTA15	85,13	923	0,032
BTA16	81,54	851	0,035
BTA17	74,83	797	0,036
BTA18	65,90	872	0,032
BTA19	63,97	868	0,029
BTA20	71,89	871	0,046
BTA21	71,41	814	0,044
BTA22	61,32	666	0,032

BTA23	54,40	674	0,040
BTA24	62,61	672	0,040
BTA25	42,72	566	0,035
BTA26	51,52	584	0,037
BTA27	45,34	499	0,034
BTA28	46,10	541	0,036
BTA29	51,37	621	0,038

Tabela 2. Informação estatística da media r² em diferentes pares de distancias entre SNP com painel de 33k em população da raça Gir selecionada para crescimento pós desmama

Cromossomo	25-50 Kb	50-100 Kb	100-500 Kb	0,5-1 Mb
BTA1	0,312±0,355	0,272±0,332	0,156±0,225	0,026±0,052
BTA2	0,222±0,267	0,200±0,246	0,146±0,188	0,025±0,050
BTA3	0,369±0,354	0,275±0,325	0,139±0,222	0,026±0,056
BTA4	0,290±0,311	0,254±0,294	0,154±0,207	0,027±0,052
BTA5	0,405±0,374	0,336±0,345	0,208±0,253	0,032±0,067
BTA6	0,288±0,320	0,249±0,287	0,176±0,223	0,029±0,061
BTA7	0,322±0,343	0,262±0,307	0,157±0,211	0,029±0,057
BTA8	0,279±0,329	0,249±0,298	0,155±0,211	0,031±0,060
BTA9	0,337±0,377	0,289±0,346	0,192±0,260	0,033±0,069
BTA10	0,302±0,341	0,264±0,300	0,155±0,208	0,028±0,059
BTA11	0,305±0,321	0,237±0,285	0,142±0,193	0,028±0,058
BTA12	0,283±0,326	0,236±0,290	0,153±0,206	0,035±0,065
BTA13	0,268±0,307	0,206±0,249	0,155±0,207	0,031±0,064
BTA14	0,291±0,317	0,262±0,294	0,156±0,208	0,032±0,062
BTA15	0,366±0,354	0,285±0,310	0,181±0,243	0,031±0,067
BTA16	0,282±0,331	0,245±0,304	0,171±0,232	0,034±0,072
BTA17	0,272±0,312	0,237±0,292	0,173±0,223	0,035±0,070
BTA18	0,153±0,316	0,243±0,286	0,165±0,223	0,032±0,066
BTA19	0,283±0,345	0,229±0,291	0,163±0,239	0,029±0,067
BTA20	0,435±0,355	0,410±0,334	0,299±0,300	0,046±0,096
BTA21	0,331±0,327	0,275±0,321	0,223±0,279	0,044±0,102
BTA22	0,243±0,261	0,208±0,247	0,127±0,174	0,031±0,059
BTA23	0,332±0,366	0,293±0,349	0,178±0,242	0,039±0,080
BTA24	0,216±0,325	0,253±0,280	0,174±0,222	0,040±0,081
BTA25	0,248±0,299	0,201±0,259	0,132±0,185	0,034±0,066
BTA26	0,206±0,222	0,198±0,227	0,149±0,190	0,037±0,073
BTA27	0,297±0,320	0,232±0,273	0,137±0,188	0,034±0,064
BTA28	0,228±0,243	0,219±0,245	0,151±0,191	0,036±0,070
BTA29	0,269±0,322	0,233±0,279	0,139±0,191	0,037±0,067

Com aumento das distâncias, houve redução do r^2 , e a média de 100 a 500kb e 0,5 a 1Mb foi maior para BTA20, BTA21 (0,299, 0,223 e 0,046, 0,044 respectivamente) e menor para BTA22 e BTA2 (0,127 e 0,025), Nas distâncias depois de 1Mb, o r^2 foi semelhante para todos os SNPs.

Os níveis de desequilíbrio de ligação ate a distancia 100kb (Figura 1), mostram que o DL diminui com o aumento das distancias, sendo útil ate a distancia de 100kb. Este decaimento pode ser causado por ser animais zebuínos com genótipos mas heterozigotos.

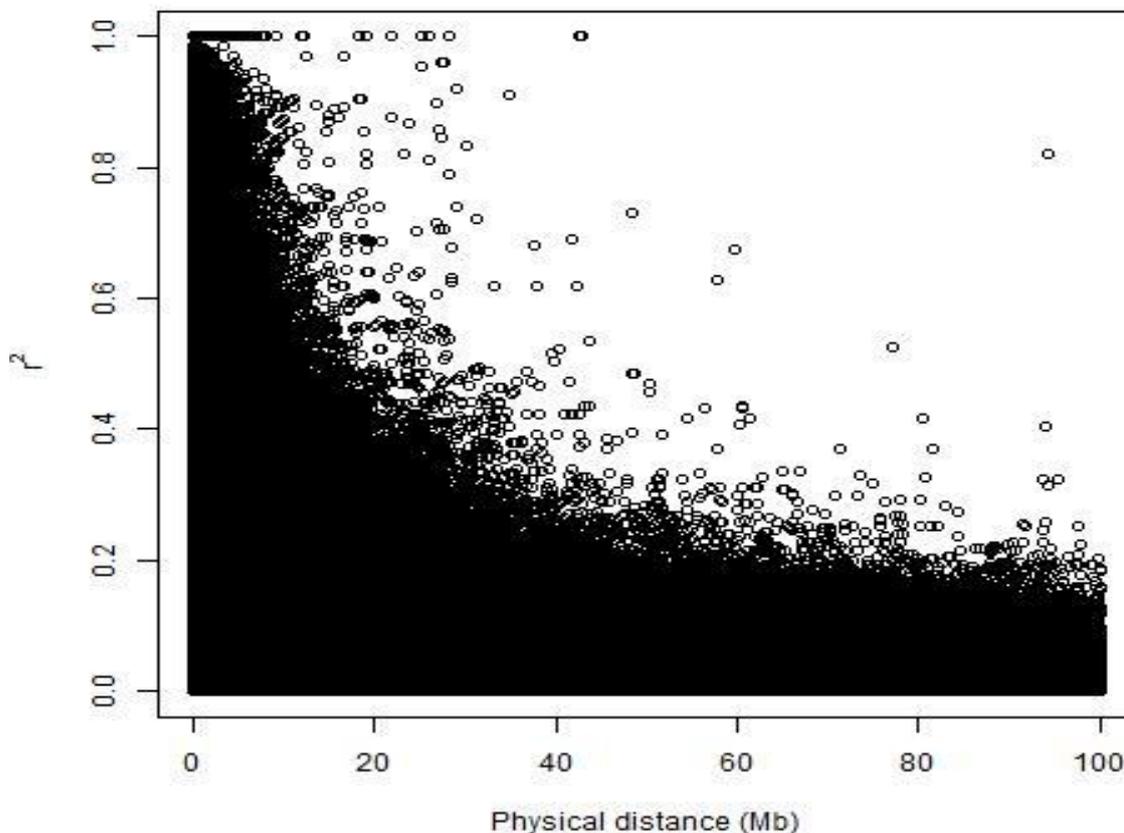


Figura 1. Níveis de desequilíbrio de ligação (DL) nos cromossomos autossômicos de população da raça Gir selecionada para crescimento pós desmama.

Na tabela 3 mostram-se medias de DL do painel 33kb para as distâncias de 25 a 50kb, 50 a 100kb, 100 a 500kb e 0,5 a 1Mb foram de 0,29, 0,25, 0,16 e 0,032, respectivamente. Estes Valores semelhantes foram descritos por Zhu et al, (2013), na raça Simental na distância de 25 a 50 kb, onde descreveram valor de r^2 igual a 0,23 utilizando painel de Bovine 50kb, e Laodim et al, (2015) com r^2 de 0,20 na distância de 50kb na raça leiteira Thailandesa.

Estudo em animais zebuínos, raça Nelore, para as distâncias de 0 a 30kb, o r^2 obtido foi igual a 0,20 para animais com o painel BovineHD777kb (ESPIGOLAN et al., 2013). Evidencia-se diferença do r^2 para distâncias semelhantes, Nas raças Hereford e Braford, o r^2 foi 0,20 e 0,15 nas distâncias de 40 a 60kb e 0,17 e 0,13 para as distâncias 50 a 100kb com painel de 50kb (BIEGELMEYER et al., 2016), De Ross et al, (2008) estudaram o DL nas raças Holstein-NZL e Angus-AUS em distâncias de 70 a 75kb, com média de r^2 de 0,10 e 0,24, enquanto Zhu et al, (2013) encontraram na distância 50 a 100kb, média de 0,16 para raça Simental.

Tabela 3. Médias do r^2 para cada distância no painel de 33kb de população da raça Gir selecionada para crescimento pós desmama

Distância	Número de SNPs	r^2
25-50 Kb	12,520	0,29
50-100 Kb	17,645	0,25
100-500 Kb	89,477	0,16
0,5-1 Mb	168,254	0,032

Outras pesquisas obtiveram um r^2 para as raças Guzerá, Holandesa e Gir, de 0,16, 0,17 e 0,21 em distancias cortas (BOHAMANOVA et al., 2010, SILVA et al., 2010). Valores menores foram descritos, r^2 igual a 0,11 e 0,15, em distância menor que 100kb para animais *Bos primigenius indicus* (MACKAY et al., 2007, ESPIGOLAN et al., 2013, respectivamente). A distância de 100kb para estimação do r^2 é muito utilizada na discussão dos resultados, porque é a distância onde o DL ainda é útil e em seguida, começa a decair com o aumento das distancias.

Silva et al, (2010) encontraram valores semelhantes na raça Gir para as distâncias de 0,5Mb, com r^2 igual a 0,142 e para a distância de 1Mb a média do r^2 foi 0,112 (SILVA et al., 2010).

Outros estudos encontraram resultados discrepantes em animais *Bos primigenius taurus*, na raça holandesa estimaram o DL mostram decaimento drástico depois de 60kb, 100kb, 500kb, 1Mb, onde o r^2 foi diminuindo de 0,16, 0,11, 0,08, 0,06, respectivamente. Na raça Angus, o r^2 foi igual a 0,30 nas distâncias >40kb utilizando o painel BovineHD, segundo a literatura os animais taurinos apresentam maior DL por serem animais mais “puros” (BOHMANOVA et al., 2010, PORTO-NETO et al., 2014, PÉREZ O'BRIEN et al., 2014).

Os trabalhos deste projeto foram realizados com painéis de maior densidade, evidenciando aumento do r^2 para todas as distâncias com relação ao painel de menor densidade, sendo obtido maior r^2 neste estudo em comparação aos descritos na literatura para animais zebuínos (MACKAY et al., 2007). Este aumento, sugere que segmentos longos de haplótipos são compartilhados pelos animais aparentados (BROWNING & BROWNING, 2011). Como evidenciado no estudo da raça holandesa Americana, onde estimaram o DL para haplótipos de herança materna, valor de r^2 igual a 0,8 em distancias menores a 20kb (KIM & KIRKPATRICK, 2009).

Os resultados indicam que as características específicas da população Gir avaliada possibilitaram o incremento do DL, entre as quais podem ser descritas a utilização de somente reprodutores nascidos dentro do rebanho, não ocorrendo entrada de material genético diferente, ou seja, o rebanho era fechado por mais de 30 anos e possuía uso intensivo de número pequeno de reprodutores, incrementando a endogamia e diminuindo a variabilidade genética.

Após a imputação de genótipos de 33kb a 777 HD, obteve-se 732.962 SNPs em todo o genoma, e com estes dados foi estimado também o valor de DL. Neste estudo as imputações foram realizadas com o painel de 33 kb, considerado “baixo” para ser imputado para HD, entretanto a acurácia da imputação foi alta (Tabela 4). Como alelos raros são responsáveis por grande proporção da variância genética, torna-se interessante obter acurácia de imputação para SNPs de alelos raros (frequência do alelo menor, $MAF < 0,05$) (Cirulli et al., 2010). Com a diminuição da MAF, a sensibilidade dos métodos de imputação diminui, apresentando maiores chances de erros de imputação.

A precisão da imputação com base nos resultados da taxa de concordância e IQS, além dos resultados médios da correlação, demonstram alta qualidade da imputação como um todo. O uso do programa FImpute é considerado de acurácia melhor para imputação para SNPs com baixa MAF. As características de interesse são influenciadas por variantes genéticas raras ($MAF < 0,05$) que tendem a ser recentes e serem associadas com haplótipos longos, o programa FImpute é eficiente em explorar haplótipos longos, normalmente, de parentes próximos. De acordo com Sargolzaei et al. (2014) o ganho na acurácia da imputação de variantes raras, aumenta com o tamanho da população referência e com a densidade do painel. Estes resultados sugerem que foram impulsionados pela relação de parentesco existente entre os animais, advindos de estudar um rebanho fechado, e pelos SNPs de baixa frequência, principalmente SNPs com $MAF \leq 2\%$.

Neste estudo as imputações foram realizadas com o painel de 33 kb, ao HD, a acurácia da imputação, na tabela 4 mostram os percentagens das três metodologias (taxa de concordância, score da qualidade de imputação, correlação de Pearson) com os 18 animais (matrizes e touros) genotipados em HD.

Tabela 4. Taxa de concordância (TC), índice de qualidade da imputação (IQS) e correlação simples (R^2) da imputação com Bovine HD BeadChip de animais da raça Gir selecionados para crescimento pós desmama.

Animal	TC	IQS	CS	Categoria
1240	0,88	0,92	0,94	Touro
1251	0,95	0,97	0,98	Touro
1253	0,89	0,93	0,94	Touro
1261	0,91	0,94	0,95	Touro
1263	0,92	0,95	0,96	Touro
830	0,91	0,94	0,95	Matriz
1194	0,86	0,91	0,92	Matriz
1246	0,96	0,97	0,98	Matriz
1258	0,93	0,96	0,97	Matriz
1260	0,96	0,97	0,98	Matriz
1262	0,93	0,96	0,97	Matriz
1264	0,91	0,94	0,95	Matriz
1265	0,95	0,97	0,97	Matriz
641	0,90	0,93	0,95	Matriz, MV
621	0,92	0,95	0,96	Matriz, MV
526	0,92	0,95	0,96	Matriz, MV
523	0,92	0,95	0,96	Matriz, MV
3888	0,94	0,96	0,97	Animal jovem

No gráfico 1 se observa os maiores valores de TC e IQS foram entre 96 e 97 onde cinco matrizes e um touro têm maiores valores em comparação com os demais animais, mostrando que os genótipos das matrizes foram, mas acurados que os genótipos dos touros segundo Johnson et al,(2013) a imputação com base na taxa de concordância com mais do 30% do total de SNPs imputados, alcança uma boa qualidade da imputação. Em estudo com imputação na raça Gir utilizando painéis de diferentes densidades, foi descrito que a acurácia apresenta maior valor quando se utiliza a informação de família, e que a precisão do programa FImpute é confiável (BOISON et al., 2015).

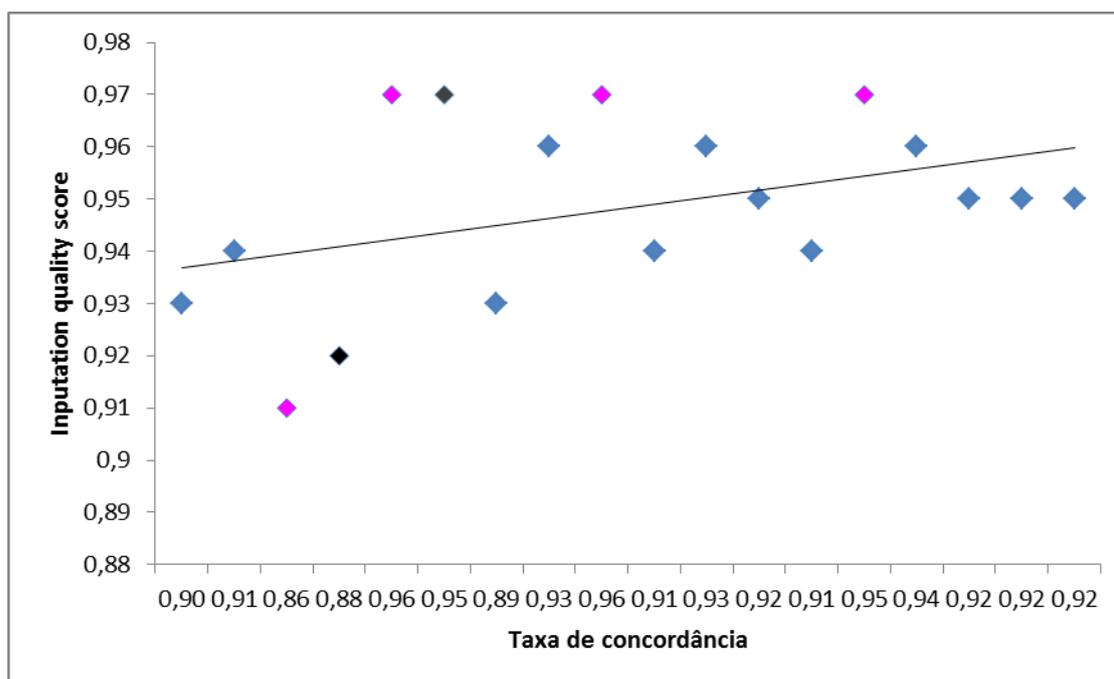


Grafico 1. Porcentagem da Taxa de concordância (TC) e Score da qualidade de imputação (IQS).

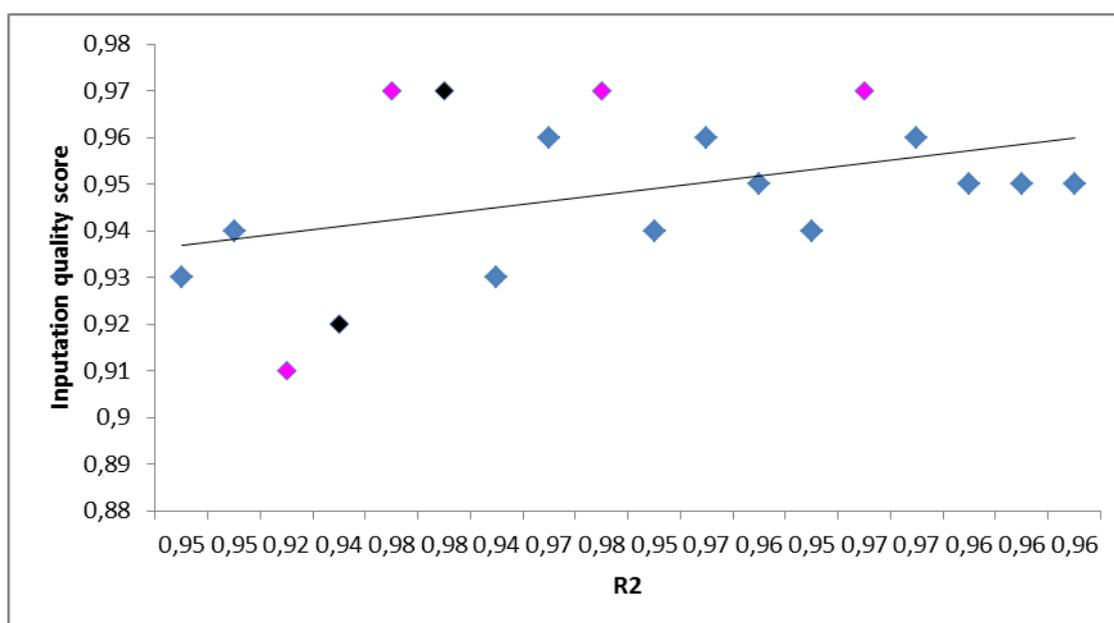


Grafico 2. porcentagem da correlação de Pearson e Score da qualidade de imputação (IQS).

Realizaram-se as mesmas análises de r^2 para os dados imputados, A Tabela 4 mostra as diferentes distâncias de 25 a 50kb, 50 a 100kb, 100 a 500kb e 0,5 a 1Mb, sendo as médias iguais a 0,35, 0,29, 0,16 e 0,032, respectivamente, obtendo estimativas maiores ao obtido no painel de 33kb.

O DL foi significativamente diferente entre os cromossomos, Obtendo média de r^2 para os SNPs separados por intervalos de 25 a 50 kb, 50 a 100 kb, 100 a 500 kb, 0,5 a 1 Mb em cada cromossomo autossômico (Tabela 4). O valor médio de r^2 para distâncias de 25 a 50kb nos diferentes cromossomos foram altos para BTA9 e BTA21, sendo 0,387 e 0,393 respectivamente, e baixos para BTA27 (0,303). A média de r^2 na distância 50 a 100kb foi maior para BTA9, BTA21 (0,335 e 0,346) e baixo para o BTA29 com 0,266.

Com aumento das distâncias, houve redução do r^2 , na média de 100 a 500kb os maiores valores foram para BTA9 e BTA21 (0,213 e 0,234, respectivamente) e menor para BTA19 (0,160). Para 0,5 a 1Mb foi maior para BTA20 e BTA23 (0,040 e 0,046 respectivamente) e menor para BTA2 (0,017). Nas distâncias depois de 1Mb, o r^2 foi semelhante para todos os SNPs.

Tabela 4. Informação estatística da media r2 em diferentes pares de distâncias entre SNP com painel imputado HD na população da raça Gir selecionada para crescimento pós desmama.

Cromossomo	25 - 50kb	50 -100kb	100 - 500Kb	0,5 -1Mb
BTA1	0,357±0,362	0,305±0,336	0,189±0,254	0,024±0,053
BTA2	0,355±0,356	0,304±0,331	0,192±0,253	0,017±0,048
BTA3	0,348±0,358	0,293±0,330	0,177±0,247	0,024±0,052
BTA4	0,344±0,354	0,289±0,326	0,179±0,246	0,025±0,053
BTA5	0,386±0,371	0,329±0,347	0,208±0,267	0,030±0,065
BTA6	0,380±0,364	0,323±0,340	0,212±0,269	0,028±0,062
BTA7	0,367±0,363	0,312±0,337	0,192±0,257	0,029±0,061
BTA8	0,344±0,350	0,292±0,323	0,182±0,245	0,031±0,062
BTA9	0,387±0,365	0,335±0,343	0,213±0,265	0,028±0,060
BTA10	0,341±0,360	0,287±0,333	0,182±0,255	0,026±0,062
BTA11	0,358±0,365	0,301±0,338	0,184±0,254	0,027±0,061
BTA12	0,336±0,353	0,285±0,326	0,183±0,251	0,033±0,066
BTA13	0,372±0,363	0,316±0,338	0,202±0,258	0,032±0,070
BTA14	0,341±0,348	0,288±0,321	0,178±0,242	0,030±0,061
BTA15	0,337±0,353	0,285±0,325	0,180±0,244	0,029±0,063
BTA16	0,358±0,360	0,301±0,332	0,194±0,256	0,032±0,070
BTA17	0,366±0,361	0,308±0,335	0,191±0,253	0,033±0,068
BTA18	0,343±0,359	0,283±0,327	0,173±0,240	0,031±0,064
BTA19	0,313±0,347	0,260±0,316	0,160±0,234	0,027±0,060
BTA20	0,359±0,358	0,309±0,334	0,204±0,263	0,040±0,080
BTA21	0,393±0,375	0,346±0,356	0,234±0,291	0,038±0,089
BTA22	0,327±0,348	0,272±0,318	0,165±0,233	0,032±0,065
BTA23	0,358±0,355	0,310±0,332	0,212±0,267	0,046±0,092
BTA24	0,373±0,367	0,319±0,342	0,201±0,264	0,038±0,082
BTA25	0,338±0,358	0,278±0,328	0,162±0,239	0,033±0,070
BTA26	0,330±0,355	0,272±0,323	0,165±0,238	0,035±0,073
BTA27	0,303±0,335	0,252±0,304	0,157±0,223	0,033±0,066
BTA28	0,322±0,340	0,271±0,310	0,173±0,233	0,036±0,075
BTA29	0,318±0,346	0,266±0,316	0,164±0,234	0,035±0,069

O painel de 777kb apresentou maiores medias de r2, em comparação com o painel de 33kb (Tabela 5). Espigolan et al, (2013) descreveram valores menores nas distâncias de 30 kb, 50 kb, 100k com r2 de 0,20, 0,16 e 0,1 respectivamente,

analisado com painel HD em animais da raça Nelore, evidenciando que o DL decai com aumento da distância, inclusive com painel de alta densidade.

Tabela 5. Estimativas médias das distâncias com dados imputados HD de população da raça Gir selecionada para crescimento pós desmama.

Distância	Número de SNPs	r^2
25-50kb	15,265	0,35
50-100kb	23,773	0,29
100-500kb	112,432	0,18
0,5-1Mb	215,643	0,032

Em estudo realizado com as raças Brahman e Angus foi descrito r^2 igual a 0,25 e 0,46, na distância de 10 kb, demonstrando que o DL é alto somente em distâncias curtas e decai quando a distância aumenta (PORTO-NETO et al., 2014). De Roos et al, (2008), estimaram o DL nas raças Holandesa e Angus obtendo r^2 com valor de 0,50 e 0,60 na distância de 5 kb, na distância de 75 kb, o r^2 foi igual a 0,10 e 0,24 respectivamente, evidenciando que o DL diminuí com o aumento das distâncias também para animais taurinos.

Segundo a literatura, os valores de DL são maiores para animais taurinos, indicando que são animais mais homozigotos, e conseqüentemente com DL inferior aos zebuínos, Estes por sua vez, são mais heterozigotos, com maior variabilidade na formação das raças, e como demonstrado em estudos na literatura, a diminuição do DL tende a ser mais rápida nas raças zebuínas em comparação com as taurinas. Por isto é recomendável a utilização de painéis de maiores densidades nos zebuínos (PÉREZ O'BRIEN et al., 2014; MACKAY et al., 2007).

Neste estudo, o DL foi maior em comparação com os trabalhos discutidos, evidenciando persistência nas distâncias até 100 kb e somente apresentando diminuição a partir de 500 kb. Em comparação com o painel de 33 kb, os valores foram inferiores em todas as distâncias, mas o DL persistiu também até 100 kb, decaindo em 500 kb. Estes resultados, indicam que os valores apresentam diferenças com os outros estudos devido a estrutura populacional dos animais avaliados da raça Gir.

O valor do DL não depende somente do painel de alta densidade, também é importante o tamanho e estrutura da população, por que ocorre influência do valor do DL e do N_e , Quando os genótipos são imputados é importante escolher os

painéis, o programa para imputar e o tipo de população, Segundo Chud et al, (2015), os painéis com menos de 50 mil marcadores não são adequados, para fazer a imputação para HD, e não permitem obter precisão aceitável na realização de análises genômicas.

A Seleção Genômica depende do desequilíbrio de ligação, que tem em determinado gene para uma característica, quando os animais são aparentados se espera que o DL seja maior para a característica e pode variar para cada população e raças Isto é provavelmente devido a diferenças no desequilíbrio de ligação entre SNPs (HAYES, LEWIN e GODDARD, 2013). Um estudo da precisão DL na seleção genômica da raça holandesa para característica de rendimento de proteína foi de 0,48, e 0,29 para o rendimento de gordura (HABIER et al., 2010). Isto evidencia que o DL varia, além disso, ao selecionar em GEBVs a extensão de LD entre SNPs e QTL, determina a fixação de alelos QTL (DEKKERS et al., 2009) e, assim, uma possível precisão das previsões genômica dos SNPs é delimitada pela proporção da variância genética do SNP, que é por sua vez é determinado pelo desequilíbrio de ligação entre o SNP.

Endogamia genômica (FROH)

Valor da endogamia genômica F_{ROH} , é estimado a partir de ROH que são longos segmentos de autozigosidade (PURCELL et al., 2007). O F_{ROH} foi de 32%, mostrando que a população de Gír tem alto nível endogamia e apresentam genótipos homozigotos, Provavelmente a explicação para este valor de F_{ROH} é resultado de um pequeno tamanho efetivo populacional (N_e) e alto desequilíbrio de ligação. A estimação do F_{ROH} em populações com elevado desequilíbrio de ligação (LD) pode encontrar endogamia recente e antiga, proporcionando uma boa estimativa de consanguinidade para identificar ROH curto com SNPs que são idênticos por descendência (MARRAS et al., 2013).

Em vários estudos de comparação entre a endogamia com pedigree (F_{PED}) e o F_{ROH} , mostram que o F_{PED} subestima o valor verdadeiro da endogamia, sendo mais confiável a estimação de endogamia com dados genômicos (SCRAGGS et al., 2013; ŠIDLOVÁ et al., 2015; GURGUL et al., 2016). Na literatura não encontra-se valores semelhantes aos obtidos neste estudo, mais Kim et al, (2015) encontrou

um valor de F_{ROH} que variaram de 0,02 para endogamia antiga e 29% para endogamia recente na raça Jersey. A variação do F_{ROH} deste estudo provavelmente deve-se ao fato de que população do Gir do IZ é caracterizada por ser um rebanho fechado por mais de trinta nos, sem uso de reprodutores de fora do programa nos acasalamentos, afetando os parâmetros populacionais. Com isso o ROH pode rastrear e informar sobre a história da população, efeitos dos gargalos e pressão de seleção, (PURFIELD et al., 2012).

Tamanho efetivo da população

A estimativa do tamanho efetivo da população (N_e) Gir analisada, com 20 gerações anteriores é apresentada na Figura 4, observa-se que o N_e foi maior na geração 20, com 30 animais, e decaiu drasticamente a partir da quinta (5) geração, sendo de 9 animais a última geração, tamanho pequeno para uma população.

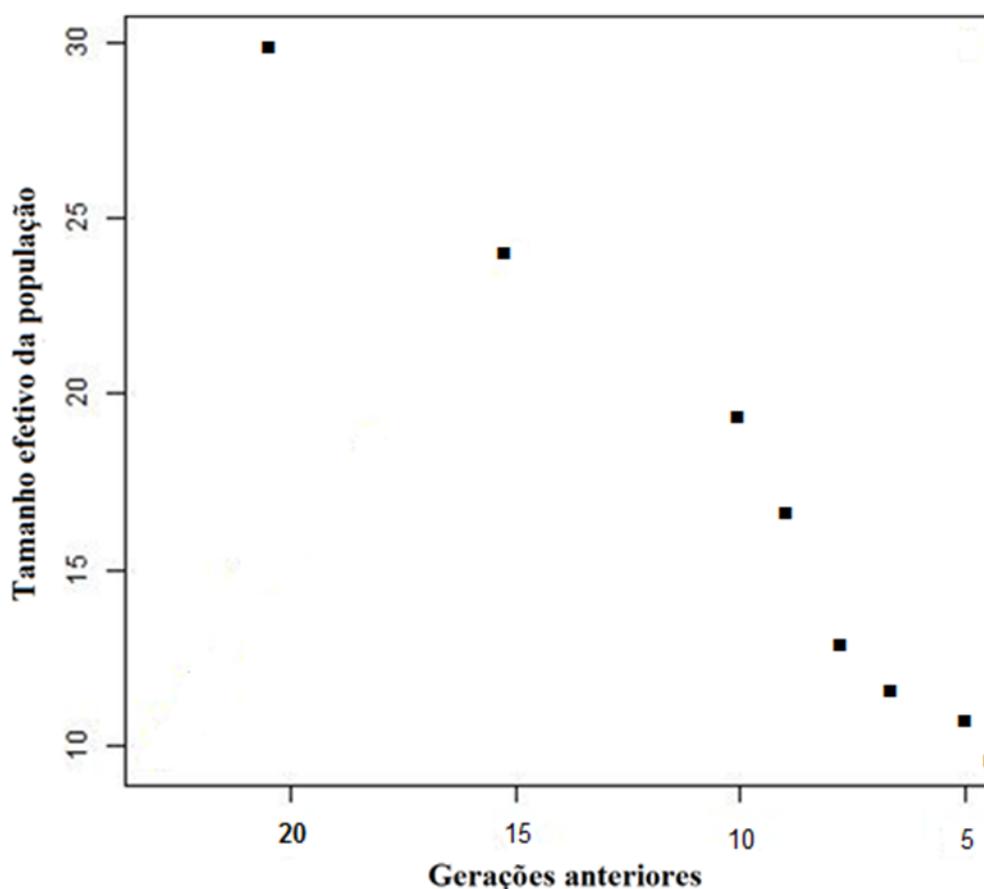


Figura 2. Tamanho efetivo de população da raça Gir selecionada para crescimento pós desmama.

Na literatura os trabalhos com Ne apresentam valores maiores, devido ao tamanho das populações e ao DL, Biegelmeyer et al., (2016) estudando populações das raças Hereford e Braford obtiveram Ne na última geração igual a 153 e 220, respectivamente. Em estudo com cavalos puro-sangue, o Ne foi igual a 190 nas duas últimas gerações (CORBIN et al., 2010). Segundo Uimari et al, (2011), em raças suínas Landrace e Yorkshire, o Ne foi de 80 e 55, apresentando-se como valores pequenos, sugerindo como causa o aumento da endogamia. Os estudos demonstram que o Ne varia de acordo com a espécie, tipo de produção e tamanho da população.

Investigações com resultados similares foram encontrados, os autores demonstraram o declínio do tamanho efetivo, na população da raça Hanwoo, com Ne igual a 98,1 nas três gerações anteriores, em animais da raça Simental, com Ne igual a 73, nas cinco gerações anteriores e em animais da raça Gir, com Ne igual a 54 e 39, nas três e duas gerações anteriores (LEE et al., 2011; ZHU et al., 2013; SILVA et al., 2010, respectivamente).

O Ne também pode ser estimado com a informação do pedigree, Faria et al, (2001) apresentando trabalho da estrutura da população da raça Gir no Brasil, estimou o Ne com 70 animais na última geração no ano de 1998.

Scraggs et al. (2013) afirmaram que o valor calculado com base nos registros genealógicos pode superestimar o verdadeiro Ne, Sugerindo que o Ne diminui em populações com alta pressão de seleção, caso dos dados avaliados neste estudo, em que o valor do Ne foi baixo e o DL foi alto, evidenciando acasalamentos entre parentes, maior endogamia, e perda da variabilidade genética. Quando o Ne é pequena, o esperado é ver um número relativamente pequeno de haplótipos segregantes na população mesmo examinaram muitos loci dentro do mesmo segmento (MACLEOD , 2010).

Os fatores relacionados como atuantes na diminuição do Ne devem ter controle rigoroso, devido ao aumento da endogamia e perda de deriva genética, que ocasionam o efeito gargalo, devido à redução drástica do tamanho efetivo da população, onde genótipos são impedidos de participar na geração seguinte ocasionando a diminuição da variabilidade genética (ARAKI & TACHIDA, 1997).

Estudo adicional da população foi realizada, onde se estimaram os parâmetros populacionais utilizando o programa ENDOG (GUTIÉRREZ &

GOYACHE, 2005), e foi obtido o coeficiente médio de endogamia (F), coeficiente médio de parentesco (AR), tamanho efetivo da população (Ne) e o número efetivo (fe, fa e fg), os valores foram iguais a 5,46%; 7,18%, 44, 23, 22 e 21, respectivamente,

A estrutura populacional do rebanho indica intensa utilização de pequeno número de reprodutores, incrementando a endogamia e afetando a probabilidade de origem do gene, Indicando que os parâmetros populacionais estão relacionados com o manejo do sistema de acasalamento, com elevado impacto na variabilidade genética,

Parâmetros populacionais

As informações de endogamia (F) e parentesco aditivo (AR) da tabela 6 possuem incremento nos valores encontrados, perdendo variabilidade genética por meio do uso dos mesmos touros nos acasalamentos, Segundo Oliveira et al, (2011) o valor de F de 3,02% não pode ser considerado elevado em rebanhos fechados para raça Nelore, porém, deve-se ocorrer melhor manejo da seleção dos animais na população recente,

Tabela 6 -Valores de Coeficientes de endogamia (F) e parentesco (AR), Tamanho efetivo da População por aumento individual de endogamia (Ne),

	População total
AR – Coeficiente Parentesco (%)	6,24
F – Coeficiente médio de endogamia(%)	7,18
Ne – Tamanho efetivo da População por aumento individual de endogamia (n)	44,17

O tamanho efetivo da população por aumento individual de endogamia (Ne) foi de 44, onde pode-se evidenciar o aumento na população , devido a causa da endogamia e o coeficiente de parentesco que foi maior na população, Segundo a FAO o valor mínimo de Ne se aproxima de 50, no entanto, deve-se considerar a importância do estudo individual de cada população, pois o Ne pode ser influenciado por meio do coeficiente de endogamia e parentesco, Em estudo desenvolvido por Dias (2012) em bovinos Caracu o Ne por aumento individual de

endogamia foi de 62, O decréscimo drástico do valor de N_e , pode representar a perda de variabilidade e impedir a formação das próximas progênes.

Na tabela 7 o número de ancestrais que contribuíram para a população foram de cinco animais, que explicam 50% da variabilidade genética, O número efetivo de fundadores (f_e), ancestrais (f_a) e genomas fundadores (f_g) foram de 23, 22 e 21, Esses resultados indicam que a diminuição dos fundadores e ancestrais no grupo foi devida a perda da variabilidade genética, possivelmente por acasalamentos entre parentes, Valores semelhantes foram encontrados por Oliveira et al, (2011) para a raça de bovinos Nelore, da linhagem Lemgruber, pertencente a rebanho também fechado, onde f_e e f_a foram 25 e 21.

Tabela 7 Estimativa da probabilidade de origem do gene e variabilidade genética de animais da raça Gir do IZ

	População total
Número total de animais	3,826
Número de fundadores	475
Número efetivo de fundadores (f_e)	23
Número efetivo de ancestrais (f_a)	22
Número efetivo de genomas fundadores	21
Razão f_e/f_a	1,05
NºA_50% *	7
NºA_100% **	162

A razão entre o número efetivo de fundadores e ancestrais (f_e/f_a) foi de 1,29, Valores próximos de 1, é considerado como valor ideal, evidenciando início do efeito gargalo (perda de variabilidade genética) causado pela diminuição da deriva genética, mudanças no tamanho da população com presença de efeito gargalo pode aumentar o LD.

Em estudo realizado por REIS FILHO (2006) com a raça Gir foi observado razão de 1,95, e no estudo de FARIA et al, (2001) na raça Nelore mocho foi obtido o valor de 1,44, Estes trabalhos indicam uso intenso de poucos reprodutores, incrementando a endogamia e afetando a probabilidade de origem do gene, Os parâmetros populacionais estão relacionados com o manejo do sistema de acasalamento de

cada população, devido à presença de elevado impacto na variabilidade genética, Com estes estudos se demonstram que os parâmetros populacionais explicam o aumento do DL e o baixo tamanho efetivo genômico, tendo uma semelhança com os resultados obtidos com informação do pedigree da população de Gír.

Dadas as características desta população com alto DL e baixo Ne a acurácia é mais precisa por ser animais com alelos idênticos por descendência.

Conclusões

Em conclusão, estes parâmetros primários de genética molecular, desequilíbrio de ligação e tamanho efetivo genômico, foram importantes para averiguar e esclarecer a relação entre os SNPs nesta população de bovinos Gir do IZ. Como o DL foi caracterizado como alto, esta população fechada pode ser importante fonte de pesquisa em estudos de GWAS e seleção genômica em bovinos de corte. O tamanho efetivo genômico foi alarmante em todas as gerações, sugerindo perda da variabilidade genética como consequência da seleção intensiva provavelmente oriundo do uso de pequeno número de reprodutores. Para incrementar a variabilidade genética e aumentar o tamanho efetivo desta população, a introdução de outros linhagens de bovinos da raça Gir deveria ser preconizada pelos melhoristas.

Referências bibliográficas

ARAKI, H.; TACHIDA, H. Bottleneck effect on evolutionary rate in then early neutral mutation model, *Genetics*, 147, 907–914, 1997.

BIEGELMEYER, P; GULIAS-GOMES, C.C.; CAETANO, A.R. Linkage disequilibrium, persistence of phase and effective population size estimates in Hereford and Braford cattle, *BMC Genetics*, 17, 1-12, 2016.

BOHMANOVA, J.; SARGOLZAEI, M.; SCHENKEL, F. Characteristics of linkage disequilibrium in North, *BMC Genomics*,, 11:421, 2010.

BOISON, S.A.; SANTOS, D.J.A.; UTSUNOMIYA, A.H.T. Strategies for single nucleotide polymorphism genotyping to enhance genotype imputation in Gyr (*Bos indicus*) dairy cattle: Comparison of commercially available single nucleotide polymorphism chips, *Journal of Dairy Science*, 98, 4969-4989, 2015.

BROWNING, S.R.; ROWNING, B.L. Haplotype phasing: existing methods and new developments, *Nature Reviews Genetics*, 12, 703-714, 2011.

CHARLES, B.A. ; SHRINER, D. ; DOUMATEY, A. A genome wide association study of serum uric acid in African Americans, *BMC Med Genomics*, 4, 17, 2011.

CHUD, T.C.S.; VENTURA, R.V.; SCHENKEL, F.S. Strategies for genotype imputation in composite beef cattle, *BMC Genetics*, 16, 99, 2015.

CIRULLI E.T, GOLDSTEIN D.B. Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nat Rev Genet* 11: 415–425, 2010.

CORBIN, L.J.; BLOTT, S.C.; SWINBURNE, J.E. Linkage disequilibrium and historical effective population size in the Thoroughbred horse, *Anim Genet*, 41, 8-15, 2010.

DEKKERS J.C.M.; ZHAO, H.; YOUNG, J.M.; HABIÉR, D. Response and inbreeding from genomic selection, 13th Quantitative Trait Locus and Marker Assisted Selection Workshop, Wageningen, The Netherlands, 2009.

DE ROOS, A.P.W.; HAYES, B.J.; SPELMAN, R.J. Linkage disequilibrium and persistence of phase in Holstein–Friesian, Jersey and Angus cattle, *Genetics*, 179, 1503–1512, 2008.

DIAS, A.C. Análise da estrutura genética da raça Caracu, 2012, 39p, Dissertação (Mestrado) - Universidade Federal de Mato grosso do Sul.

ESPIGOLAN, R.; BADLI, F.; BOLIGON, A.A. Study of whole genome linkage disequilibrium in Nellore cattle, *BMC Genomics*, 14, 305, 2013.

FARIA, F.J.C.; FILHO, A.E.; MADALENA, F.E.; JOSAHKIAN, L.A. Parâmetros Populacionais do Rebanho Gir Mocho Registrado no Brasil, Rev, Bras, Zootec, 30, 1984-1988, 2001.

FALCONER, D.S.; MACKAY, T.F.C. Introduction to quantitative genetics. 4.ed. Edinburgh : Longman Group Limited. 464p, 1996.

FISHER R,A, The genetical theory of natural selection, Oxford University Press, 1930, p,308,

GODDARD, M.E ; HAYES, B.J. Genome-Wide Association Studies and Linkage Disequilibrium in Cattle, Bovine Genomics, Chapter 13, 194-195, 2012.

GURGUL, A.; SZMATOŁA, T.; TOPOLSKII, P.; JASIELCZUK, I.; ŻUKOWSKI, K.; BUGNO-PONIEWIERSKA, M. The use of runs of homozygosity for estimation of recent inbreeding in Holstein cattle, Journal of Applied Genetics, pp 527–530, 2016.

GUTIÉRREZ, J.P.; GOYACHE, F. A note on ENDOG: a computer program for analyzing pedigree information, Journal of Animal Breeding and Genetics, 122, 172 -176, 2005.

HABIER, D.; TETENS, J.; SEEFRIED, F.R.; LICHTNER, P.; THALLER, G. The impact of genetic relationship information on genomic breeding values in German Holstein cattle, Genetics Selection Evolution 42:5, 2010.

HAYES, B.J. ; VISSCHER, P.M. ; MCPARTLAN, H.C. Novel Multilocus Measure of Linkage Disequilibrium to Estimate Past Effective Population Size, Genome researche, 13, 635-643, 2003.

HAYES, B.J.; LEWIN H.A.; GODDARD, M.E. The future of livestock breeding: genomic selection for efficiency, reduced emissions intensity, and adaptation, Trends in Genetics, 2013.

HUANG, L. ; LI, Y. ; SINGLETON, A.B. Genotype imputation accuracy across worldwide human populations, *Am J Hum Genet*, 84, 235–250, 2009.

JOHNSON, E.O.; HANCOCK, D.B.; LEVY, J.L. Imputation across genotyping arrays for genome-wide association studies: assessment of bias and a correction strategy, *Hum Genet*, 132, 509-522, 2013.

KIM, E.S.; SONSTEGARD, T.S.; VAN TASSELL, C.P.; WIGGANS, G.; ROTHSCHILD, M.F. The Relationship between Runs of Homozygosity and Inbreeding in Jersey Cattle under Selection, *PLoS One*, 10(7): e0129967, 2015.

LAODIM, T.; KOONAWOOTRITTRIRON, S.; ELZO, M.A. Genome-wide linkage disequilibrium in a Thai multibreed dairy cattle population, 180, 27-33, 2015.

LEE, S.H.; CHO, Y.M.; LIM, D. Linkage Disequilibrium and Effective Population Size in Hanwoo Korean Cattle, *Asian-Aust, J, Anim, Sci*, 24, 1660-1665, 2011.

LIN, P.; HARTZ, S.M. ; ZHANG, Z.; et al, A new statistic to evaluate imputation reliability, *PLoS One*, 5, e9697, 2010.

MACKAY, S.; SCHNABEL, R.; MURDOCH, B. Whole genome linkage disequilibrium maps in cattle, *BMC Genetics*, 8, 74, 2007.

MACLEOD, I.M. Whole-genome resequencing of two elite sires for the detection of haplotypes under selection in dairy cattle, *CroosMark*, 2012.

MARRAS, G.; GASPA, G.; SORBOLINI, S.; DIMAURO, C.; MARSAN, P.A.; VALENTINI, A.; WILLIAMS, J.L.; MACCIOTTA, N.P. Analysis of runs of homozygosity and their relationship with inbreeding in five cattle breeds farmed in Italy, *Animal Genetics*, 2013.

OLIVEIRA, P.S. Estrutura populacional de rebanho fechado da raça Nelore da linhagem Lemgruber, *Pesquisa Agropecuária Brasileira*, v,46, p,639-647, 2011.

PÉREZ O'BRIEN, A.M.; UTSUNOMIYA, Y.T.; MÉSZÁROS, G. Assessing signatures of selection through variation in linkage disequilibrium between taurine and indicine cattle, *Genetics, Selection, Evolution*, 46, 19, 2014.

PORTO-NETO, L.R.; KIJAS, J.W.; REVERTER, A. The extent of linkage disequilibrium in beef cattle breeds using high-density SNP genotypes, *Genet Sel Evol*, 24, 46:22, 2014.

PURCELL, S.; NEALE, B.; TODD-BROWN, K.; THOMAS, L.; FERREIRA, M.A.R.; BENDER, D.; MALLER, J.; SKLAR, P.; BAKKER, P.I.W.; DALY, M.J.; AND SHAM, P.C. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses, *Am J Hum Genet*, 81(3): 559–575, 2007.

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

RAZOOK, A.G ; FIGUEIREDO, L.A ; BONILHA NETO, L.M ; et al, Selection for yearling weight in Nelore and Guzará zebu breeds: selection applied and response in 15 years of progeny, In: WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 6, 1998.

REIS FILHO, J.C. Endogamia na raça Gir, 2006, 61p, Dissertação (Mestrado) – Universidade Federal de Viçosa, Viçosa.

SARGOLZAEI, M. ; CHESNAIS, J.P. ; SCHENKEL, F.S. A new approach for efficient genotype imputation using information from relatives, *BMC Genomics*, 15, 478, 2014.

SCRAGGS, E.; ZANELLA, R.; WOJTOWICZ, A. Estimation of inbreeding and effective population size of full blood wagyu cattle registered with the American Wagyu Cattle Association, *J, Anim, Breed, Genet*, 131, 3-10, 2013.

SILVA, C.R.; NEVES, H.H.R.; QUEIROZ, S.A. Extent of linkage disequilibrium in Brazilian Gyr dairy cattle based on genotypes of AI sires for dense SNP markers, In: 9th World Congress on Genetics Applied to Livestock Production, Leipzig,

Alemanha, Proceedings of the 9th World Congress on Genetics Applied to Livestock Production, 2010.

ŠIDLOVÁ, V.; KASARDA, R.; MORAVČÍKOVÁ, N., TRAKOVICKÁ, A.; CURIK, I.; FERENČAKOVIĆ, M. genomic variability among cattle populations based on runs of homozygosity, Poljoprivreda 44-47, 2015.

SVED, J.A. Linkage disequilibrium and homozygosity of chromosome segments in finite populations, Theor, Popul, Biol, 2, 125-41, 1971.

UIMARI, P.; TAPIO, M. Extent of linkage disequilibrium and effective population size in Finnish Landrace and Finnish Yorkshire pig breeds, J Anim Sci, 3, 609-614, 2011.

WRIGHT, S. Evolution in Mendelian populations, Genetics, 16, 97-159, 1931.

ZHAO, H.; NETTLETON, D.; DEKKERS, J.C.M. Evaluation of linkage disequilibrium measures between multi-allelic markers as predictors of linkage disequilibrium between single nucleotide polymorphisms, Genetical Research, 89, 1-6, 2007.

ZHU, M.; ZHU, B.; WANG, Y.H. Linkage Disequilibrium Estimation of Chinese Beef Simmental Cattle Using High-density SNP Panels, Asian-Australasian Journal of Animal Sciences, 26, 772-779, 2013.