UNIVERSIDADE FEDERAL DE VIÇOSA

PEDIGREE RECONSTRUCTION, RELATIONSHIP AND POPULATION STRUCTURE USING SNP MARKERS IN GIR CATTLE

Arielly Oliveira Garcia
Magister Scientiae

VIÇOSA – MINAS GERAIS
2021

ARIELLY OLIVEIRA GARCIA

PEDIGREE RECONSTRUCTION, RELATIONSHIP AND POPULATION STRUCTURE USING SNP MARKERS IN GIR CATTLE

Dissertation submitted to the Animal Science Graduate Program of the Universidade Federal de Viçosa in partial fulfillment of the requirements for the degree of *Magister Scientiae*.

Adviser: Simone Eliza Facioni Guimarães

Co-advisers: João Cláudio do Carmo Panetto
Pamela Itajara Otto
Marcos Soares Lopes

VIÇOSA – MINAS GERAIS
2021

Ficha catalográfica elaborada pela Biblioteca Central da Universidade Federal de Viçosa - Campus Viçosa

T

G216p
2021

Garcia, Arielly Oliveira, 1994-
Pedigree reconstruction, relationship and population structure using SNP markers in Gir cattle / Arielly Oliveira Garcia. – Viçosa, MG, 2021.
109 f. : il. (algumas color.) ; 29 cm.

Inclui apêndices.
Orientador: Simone Eliza Facioni Guimarães.
Dissertação (mestrado) - Universidade Federal de Viçosa.
Inclui bibliografia.

1. Gir (Bovino). 2. Funções Verossimilhança. 3. Paternidade. 4. Animais - População. I. Universidade Federal de Viçosa. Departamento de Zootecnia. Programa de Pós-Graduação em Zootecnia. II. Título.

CDD 22. ed. 636.28

Bibliotecário(a) responsável: Renata de Fatima Alves CRB6/2578

ARIELLY OLIVEIRA GARCIA

PEDIGREE RECONSTRUCTION, RELATIONSHIP AND POPULATION STRUCTURE USING SNP MARKERS IN GIR CATTLE

Dissertation submitted to the Animal Science Graduate Program of the Universidade Federal de Viçosa, in partial fulfillment of the requirements for the degree of *Magister Scientiae*.

APPROVED: May 27th, 2021.

Assent:

Arielly Oliveira Garcia
Author

Simone Eliza Facioni Guimarães
Adviser

ACKNOWLEDGMENTS

Firstly, would like to thank randomness. Which, given the innumerable probabilities, gave me life and allowed me to be where I am now.

Sincere thanks to my mother Antonia M. O. Garcia and my father Adriano S. Garcia, who did and do everything to make my dreams come true. To my sister Arianny O. Garcia who since she was born teaches me about sharing, companionship, and dedication. To my dear grandmother Alzira Matos de Almeida and my grandfather Manoel José de Oliveira (*in memoriam*).

Thanks also to:

My boyfriend Vinicius P. Campos, for the presence, friendship and love in good times, and especially for support and patience in bad times;

My friends/sisters Deisinara B. Oliveira, Fernanda G. Ferreira, Larissa C. Silva, Luane O. da Silva, Jessica X. Silva and Tamiris N. S. Silva for all support given during this and other trajectories;

My pets who have always taught me, through actions, the importance of empathy, love and care. Thank you very much to those who were, those who are and those who will be. In particular, my Dinha (*in memoriam*), daughter-dog-best-friend-forever! And my Darwin son-cat-sly-forever!

All my family for the moments of joy;

My adviser, Professor Simone Eliza Facioni Guimarães for support, teachings and patience;

Universidade Federal de Viçosa (UFV) and Animal Science Department for support and the opportunity;

To Grupo de Discussão em Genética e Melhoramento Animal (GDMA) for all learning and knowledge exchange;

To the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), to granting the scholarship;

Embrapa Dairy Cattle Research Center for the partnership and for providing the data needed to carry out this study;

My co-advisor Pamela Itajara Otto for help with the analyses, incentive and advises;

My co-advisor João Cláudio do Carmo Panetto for patience and help with the analyses;

To researcher Marcos Vinicius Gualberto Barbosa da Silva for his collaboration in this study;

My friends from UFV and Animal Biotechnology Laboratory (LABTEC) for continuous knowledge, learning exchange and pleasant workplace;

All the teachers who have contributed to my training so far, thank you for never giving up! You build dreams!

“The world we live in is the one we build from our perceptions, and it is our structure that allows those perceptions. Therefore, our world is our worldview. If the reality we perceive depends on our structure - which is individual - there are as many realities as there are perceiving people.”

— Humberto Mariotti

ABSTRACT

GARCIA, Arielly Oliveira, M.Sc., Universidade Federal de Viçosa, May, 2021. **Pedigree reconstruction, relationship and population structure using SNP markers in Gir cattle.** Adviser: Simone Eliza Facioni Guimarães. Co-Advisers: João Cláudio do Carmo Panetto, Pamela Itajara Otto and Marcos Soares Lopes.

Over time estimates of genetic values and studies of the structure and genetic diversity of animal populations have been carried out using traditional methods. Using only pedigree records, it was possible to obtain valuable information on their productive potential, history of population and genetic diversity. Consequently, incorrect or incomplete pedigree information could considerably reduce the accuracy of the analyzes. Currently, the use of markers has allowed the correction of information contained in the pedigree, in addition to estimating the relationship between animals more accurately, including in the absence of a pedigree. More accurate estimates of population parameters, such as inbreeding, linkage disequilibrium and effective size are also possible. With these parameters, we can evaluate the population structure and genetic diversity, guiding management and mating strategies in herds and the population as a whole. In this context, our general objectives were i. Identify a set of at least 1000 SNP markers, present in the main commercial chips, that accurately allow parentage assignment to be carried out in Gir breed; ii. Estimate relationship coefficients in Gir cattle based on SNP markers; iii. Estimate the inbreeding coefficients in Gir animals through runs of homozygosity; iv. Estimate the effective population size based on linkage disequilibrium in Gir cattle. We used the likelihood ratio approach to reconstruct the pedigree of Gir cattle, in order to correct any errors in annotation and to deepen the information contained in the pedigree records. The approach proved to be satisfactory, improving the depth of the pedigree and showing that it is well established in terms of recent information. The relationship coefficients allowed us to assess the distribution of values for each relationship category of the reconstructed pedigree. The population structure of the Gir breed was analyzed using effective size, linkage disequilibrium and inbreeding based on runs of homozygosity (ROH). In which a decrease in effective size and levels of inbreeding were detected, especially those based on short segments of ROH, with moderate to high values, suggesting the presence of bottlenecks in the genome of Gir cattle. On the other hand, a lower percentage of ROH resided in the longest ROH classes. The present study provides important information that can be added to the Brazilian Dairy Gir Breeding

Program (PNMGL) to increase the accuracy and volume of records, in addition to assisting in the development of mating management strategies.

Keywords: Likelihood approach. Parentage analysis. Population parameters. Probability.

RESUMO

GARCIA, Arielly Oliveira, M.Sc., Universidade Federal de Viçosa, maio de 2021. **Reconstrução de pedigree, parentesco e estrutura de população com uso de marcadores SNP em gado Gir.** Orientadora: Simone Eliza Facioni Guimarães. Coorientadores: João Cláudio do Carmo Panetto, Pamela Itajara Otto e Marcos Soares Lopes.

Ao longo do tempo, estimativas de valores genéticos e estudos de estrutura e diversidade genética das populações animais têm sido realizadas por métodos tradicionais. Utilizando apenas os registros de pedigree era possível obter informações valiosas sobre o potencial produtivo, a história da população e a diversidade genética. Consequentemente, informações de pedigree incorretas ou incompletas podiam reduzir consideravelmente a acurácia das análises. Atualmente, o uso de marcadores tem permitido a correção de informações contidas no pedigree, além de estimar o parentesco entre os animais de forma mais acurada, inclusive na ausência de pedigree. Estimativas mais acuradas de parâmetros populacionais, como endogamia, desequilíbrio de ligação e tamanho efetivo também são possíveis. Com esses parâmetros podemos avaliar a estrutura populacional e a diversidade genética, orientando estratégias de manejo e acasalamento nos rebanhos e na população como um todo. Nesse contexto, nossos objetivos gerais foram: i. Identificar um conjunto de pelo menos 1000 marcadores SNP, presentes nos principais chips comerciais, que permitam com precisão a atribuição de paternidade na raça Gir; ii. Estimar coeficientes de relacionamento em gado Gir com base em marcadores SNP; iii. Estimar coeficientes de endogamia em animais Gir por meio de corridas de homozigose; 4. Estimar tamanho efetivo da população com base no desequilíbrio de ligação em gado Gir. Nós utilizamos a abordagem da razão de verossimilhança para reconstruir o pedigree do gado Gir, a fim de corrigir eventuais erros de anotação e aprofundar as informações contidas nos registros de pedigree. A abordagem provou ser satisfatória, melhorando a profundidade do pedigree e mostrando que ele é bem estabelecido quanto à informações recentes. Os coeficientes de parentesco nos permitiram avaliar a distribuição dos valores para cada categoria do pedigree reconstruído. A estrutura populacional da raça Gir foi analisada por meio do tamanho efetivo, desequilíbrio de ligação e endogamia baseada em corridas de homozigose (ROH). Em que detectamos uma diminuição no tamanho efetivo, e níveis de endogamia, especialmente aqueles baseados em segmentos curtos de ROH, com valores moderados a altos, sugerindo a presença de gargalos no desenvolvimento da população de gado Gir no Brasil. Por outro lado, uma menor porcentagem de ROH residia nas classes de

ROH mais longas. O presente estudo fornece informações importantes que podem ser agregadas ao Programa Nacional de Melhoramento do Gir Leiteiro (PNMGL) para aumentar a acurácia e o volume dos registros, além de auxiliar no desenvolvimento de estratégias de manejo de acasalamento.

Palavras-chave: Abordagem de verossimilhança. Análise de paternidade. Parâmetros populacionais. Probabilidade.

SUMMARY

Chapter 1

GENERAL INTRODUCTION

Brazil, a predominantly tropical country, is today one of the largest milk producers in the world. According to data from 2019, presented by the United Nations Food and Agriculture Organization - FAO (2021), this is in the 5th position in the ranking of world milk production. When it comes to productivity, however, the country holds an inferior position (FAO, 2021). In an attempt to improve the productivity of milk production systems under tropical conditions, zebu cattle breeds and crossbreeds are used in order to improve adaptability to the climate and management adopted, maintaining good levels of productivity, fertility and vigor (Stafuzza et al., 2017). Genetic factors are responsible for these desirable characteristics of the Zebu breeds.

Since 1952, due to experiments carried out by Hershey and Chase, there is no doubt that the DNA molecule is in fact the genetic material, and it contains the information that guides and coordinates the development of organisms, in addition to being responsible for heredity. Because it is the molecule that holds the base of all genetic information, its content mostly conserved within the same species or race. However, even among members of the same family, siblings or parents and offspring, there are small variations, called polymorphisms (Conti, Moreno & Ong, 2010). The DNA polymorphism provides genetic variability and allows the construction of an individual-specific genetic profile. DNA analysis is a powerful way to identify animals, as it remains intact even after slaughter (Harlizius, 2011).

The biotechnological and molecular genetic advances have made it possible to identify the relationship between individuals through the genotype. Obtaining true information about the degree of relationship between animals in a population is essential for genetic improvement, since genetic evaluations rely upon the relationships between individuals. One of the best known and most used designs is the progeny test. The relevance of the paternity test is due to losses in the accuracy of selection and reduction of genetic progress that can occur due to mistaken records (Van Vleck, 1970; Geldermann, Pieper & Weber, 1986). According to Ron et al. (1996), paternity errors can reach up to 20% of records, which would considerably reduce genetic progress in the population in question.

Moreover, the information from the markers can also be used in the evaluation of population parameters, such as inbreeding, linkage disequilibrium and effective population size, providing important information about the genetic architecture of the studied population.

LITERATURE REVIEW

Molecular Marker

Marker is a characteristic capable of showing differences between two or more individuals or organisms. In the area of genetics and breeding, three types of markers are considered: i. Morphological, in general phenotypic characteristics of discrete variation, such as those easy to be visually identified (coat color, presence of horn, etc.); ii. Biochemical, its use began in the 60s, are isoenzymes, molecular variants of the same enzyme that have the same or similar function in the same individual, other proteins and metabolites, and finally; and iii. Genetic molecular markers, which are capable of detecting genetic polymorphisms between two or more organisms (Nodari et al, 2016).

From a molecular point of view, a genetic marker can be in an intragenic region or in a fragment with unknown function, which is more common, and plays the role of identifying a site or a region on a chromosome. According to Nodari et al (2016), a genetic marker must have a number of attributes, such as high level of polymorphism, stability in different environments, allow identification of parents, among others. However, the simplicity and low costs of the method are determining factors for the extension of use of a molecular marker.

One of the most well-known and renowned markers worldwide is the Short Tandem Repeat (STR), also called microsatellite. STR are repetitions of one, two, three or four nucleotides, which are spread across an individual's genome; the most common in mammals being the dinucleotide $(GT)_n$ (Stallings et al., 1991). The number of repetitions varies considerably between individuals and between homologous chromosomes of the same individual, where each genomic region that contains a certain number of repetitions of one of these sequences constitutes an allele, and because it is multi-allelic it is highly informative (Nodari et al., 2016). To use microsatellites, there is a need to first amplify a region, sequence it and, finally, synthesize the specific primers for each locus. Once this is done, the marker locus can be used indefinitely in that species. Therefore, there is a high initial cost and labor.

Due to the ease and automation, which generates a considerable reduction in cost, a marker that has been spreading more and more is the Single Nucleotide Polymorphism (SNP). SNPs are elementary changes in the DNA molecule, where there is the replacement of a single nitrogenous base, caused by mutation and inherited as allelic variants (Pierce, 2017). The most common mutations occur when a base is replaced by another of the same class, a process called transition (purine ⇔ purine or pyrimidine ⇔ pyrimidine) and cross-sections are less common, when a nitrogenous base is replaced by another class (purine ⇔ pyrimidine). SNPs are characterized by sequencing DNA fragments, where there is a comparison of the strip from the sample of interest and a consensus strip referring to the species in question (Caetano, 2009). Always emphasizing that, to be considered a SNP, the allele frequency of that change in the population in question must be ≥ 1%.

Countless factors influence the choice of the marker to be used. Such as, the type of study, in which dominant or codominant nature of the marker can be crucial, laboratory facilities and especially, the costs involved. As for the comparison between markers to be used, the most important characteristics to be considered are the number of loci amplified in a single reaction, also called multiplexing capacity, the number of alleles per locus and the proportion of polymorphic loci.

More recent approaches use genetic markers, such as SNP markers, either to estimate relatedness and inbreeding coefficients with great accuracy in an unobserved pedigree or to identify shared genomic regions that are unaffected by recombination (Speed & Balding, 2014). The observed genotypes can be used to distinguish between a set of possible alternative degrees of relationship, or they can be used to estimate an unknown degree of relationship. Parentage tests are commonly done with STR markers, which are often highly accurate (Knibb et al., 2014). But, recently, genotyping using high-density SNP chips has become available at lower cost and it is increasingly being used in genetics research and breeding programs, leading to an increasing number of studies using these markers for parentage validation and assignments (Panetto et al., 2017; Huisman, 2017; Grashei, Ødegård & Meuwissen, 2018).

Parentage Analysis

Parentage analysis is a molecular technique that seeks to identify an individual's biological parent. The test is performed by collecting, analyzing and comparing the genetic material belonging to the progeny and alleged parents. The analysis is based on the Mendelian segregation law, which states that half of an individual's genetic material is inherited from the mother and the other half comes from the father. Therefore, there is a comparison of the DNA of the individuals in question in order to identify the inheritance of the genotype and give us a degree of relationship between them.

In animal breeding, to obtain genetic gain in a herd, a complete and reliable pedigree is necessary. Correct pedigree information is needed to achieve greater accuracy in predicting genetic values and estimating genetic gain, while errors in the pedigree can reduce them (Van

Vleck, 1970; Geldermann, Pieper & Weber, 1986; Senneke et al., 2004). In addition, information about genealogy is indispensable for the study of low-incidence characteristics, such as diseases, genetic defects and mortality (Harlizius, 2011)

DNA markers have been shown to be useful for verifying the information contained in the conventional pedigree, in addition to the identification of parents of individuals for whom this information were not available, which characterizes the verification and identification of parentage, respectively. Nowadays, the microsatellite is the most used marker in parentage tests, as it allows a satisfactory conclusion with the use of a few markers, which are already known and established for many different species. However, there is a huge migration trend towards the use of SNPs, since the technological advances have allowed high performance and accuracy methodologies, low cost, in addition to being used for other commercial purposes, which would dilute the costs of parentage analyzes. Other important attributes of SNPs include: i. Greater mutational stability, when compared to microsatellites; ii. Lower genotyping cost, as it is a fully automated process; iii. Considerably lower error rates since they are conducted with redundancy; and iv. The ability to easily extend in quantity, leading to the possibility of working with thousands of markers (Caetano, 2009).

The molecular information of the animals is analyzed computationally and the paternity check, based on the pedigree, is commonly done by the exclusion method, due to its ease of execution and interpretation. As for the attribution of parentage, this is accomplished through statistical analysis of the molecular data, which can be based on three methods: Exclusion, Likelihood Ratio, and Bayesian.

Exclusion Method

The basis of this method is the identification of genetic incompatibilities between two individuals, so that it can be conclusively demonstrated that one cannot be a sire, or a dam, of an offspring. The results of genotyping are subjected to statistical analysis, always taking into account the laws proposed by Mendel, in which there is segregation of alleles and the fact that half of an individual's genetic material comes from maternal inheritance and the other half from paternal inheritance (Flanagan & Jones, 2019). The exclusion power is directly related to the number of alleles and their frequency, among other factors. The probabilities of inclusion and exclusion of paternity are always complementary.

This method has the assumption that there were no mutations or errors during meiosis or in the genotyping technique, typing errors or untyped loci, an assumption whose validity varies depending on the type of locus under consideration (Araújo, 2004; Flanagan & Jones, 2019).

According to Flanagan & Jones (2019), the exclusion method, although simpler and easier to interpret, tends to become disused over time, as it uses arbitrary incompatibility values and does not take into account the probability of paternity. In addition, the exclusion approach using SNP, which are biallelic markers, uses only homozygous loci data, in which an offspring and its putative parent have different homozygous genotypes. However, a heterozygous loci also carries information on the likelihood of relationship (Marshall et al. 1998; Kalinowski, Taper & Marshall, 2007), but this information ends up being discarded in this method.

Likelihood Ratio Method

It is a method of attributing paternity based on probabilities, also called categorical attribution, where paternity is completely attributed to the supposed parent with the greatest probability. The assignment of the parents is based on the calculation and comparison of the relative probabilities of different hypotheses with respect to the relationships between the supposed dyads between parents and offspring or the sire-offspring-dam triads (Flanagan & Jones, 2019). The probability is based on the observation of the genotypes, given the hypothesis, which is the proposed relationship between individuals. The probability can then be easily calculated using the Mendelian inheritance rules (Marshall et al., 1998; Kalinowski, Taper & Marshall, 2007). In the attribution of parents, the ratio involves the hypothesis that the dyad or triad represents a true set of parents and offspring, versus the hypothesis that the individuals are not related. Generally, we deal with the logarithms of probabilities, and the likelihood ratio becomes a log-likelihood ratio or LOD score (Marshall et al., 1998). A positive LOD score indicates that the paternity hypothesis is more likely, an LOD score equal to zero implies that the alleged male is as likely a parent as any male randomly chosen from the population, while a negative LOD score indicates that the hypothesis that individuals are unrelated is more likely, given their genotypes (Marshall et al., 1998; Flanagan & Jones, 2019). These LOD values, although useful in obtaining a maximum likelihood solution, cannot be interpreted statistically or biologically.

Taylor et al. (1997) and Marshall et al. (1998) recognized that critical values for LOD scores could be determined by simulation, thus being able to attribute a level of confidence to the attribution. The approach used by Marshall et al. (1998) actually uses delta (Δ), the difference in the LOD score between the most likely parent and the second most likely parent, or the gross LOD score if only one candidate has a positive value, and simulates parent and progeny populations to determine a critical value of Δ that results in a desired confidence level and that fits the conditions of the available data. This approach was the first to control the experimental error in the paternity analysis.

Bayesian Method

Currently, the fractional allocation approach is also used for parentage analysis, known as Bayesian paternity analysis (Flanagan & Jones, 2019). Fractional allocation partially assigns offspring to parents due to *a posteriori* probabilities, where the supposed parent with the greatest *a posteriori* probability receives the largest fraction of the offspring, but the offspring would also be partially allocated to any parent with a non-zero probability. Unlike the likelihood ratio method, which can predict more than one possible parent, but still distinguish a single true parent from among these (Flanagan & Jones, 2019; Araújo, 2004).

The probability of *a posteriori* paternity is estimated by weighing the reproductive contribution of each candidate parent with the probability of paternity relative to the other candidates. The initial informative *a priori*, essential under the rules of the theorem, is that all possible parents have equal probability of paternity (Signorovitch & Nielsen, 2002).

In the beginning, this fractional approach was intuitively unpleasant, since it is strictly false from the biological perspective; fractional paternity has no biological analog. An individual cannot be 4% the parent of an offspring, for example, even if this result is possible in fractional allocations (Flanagan & Jones, 2019). This intuitive aversion led to the widespread adoption of categorical attribution, although fractional attribution has better statistical properties for estimating many values of interest (Neff, Repka & Gross, 2001; Nielsen et al., 2001).

Today, likelihood ratio method is considered the best method to be used when little information about the population and its structure is available (Flanagan & Jones, 2019).

Inbreeding and Relationship Coefficients

In animal breeding, the selection of animals to be used for reproduction is based on their genetic value. The genetic value is predicted based on information between relatives using the best linear unbiased prediction (BLUP), via equations of mixed models, as proposed by Henderson (1975), which includes the matrix of relationships between animals (Lopes, 2005). The conventional pedigree matrix obtained through manual records is generally used, which presents the genetic covariances between individuals. However, BLUP is often not used with maximum efficiency due to errors in the pedigree annotation and due to the difficulty in estimating the real relationships between individuals.

Inbreeding is a term used to specify mating between individuals more related to each other than the population average to which they belong. If two individuals have a common ancestor, it is possible that both carry replicas of one or more alleles present in the ancestor, and, when mated to each other, they can pass these replicas on to their progeny. Identical alleles from the same common ancestor are called the allele Identical-by-Descent (IBD), while identical alleles that are not inherited from the same common ancestor are called alleles Identical-by-State (IBS) (Falconer & Mckay, 1996).

According to Weir, Anderson & Hepler (2006), if we consider two unrelated individuals, their offspring will inherit exactly one allele from each locus, coming from each parent, and therefore, as they do not share an identical allele by descent (IBD), this it will give the descendant a relationship coefficient (r) equal to 0.5 with each parent. According to Visscher et al. (2002), when we evaluate full siblings they can share zero ($r = 0$), one ($r = 0.5$) or two ($r = 1$) IBD alleles for each locus. However, in practice we consider the mean value ($r = 0.5$) to quantify the relationship between complete siblings, often causing an under or overestimation of relationship that will directly interfere in the estimation of the animal's genetic value. Moreover, this value can vary considerably if there are inbreeding mating or if the IBS information is used. Speed & Balding (2015) found values ranging from 0.37 to 0.63, considering a 95% credibility range, for human half-siblings.

Before the discovery and use of SNP markers, the low availability of markers had always been described as the main limiting factor for the use of molecular markers in the estimation of inbreeding and relationship (Garant & Kruuk, 2005). Markers, such as microsatellites for example, are widely spaced, which allows us to identify the sharing of long chromosomal segments between narrower family members, but fail to detect the many minor genetic effects shared by distant relatives (VanRaden et al., 2009). Moreover, the greater the number of markers distributed throughout the genome, more accurate the measurement of genetic similarity among individuals. As these are elementary changes in the DNA molecule, the use of SNP-type markers would solve this limitation, but increasing the number of markers is also synonymous with increasing challenges, as the use of denser marker maps increases the proportion of markers strongly connected and, consequently, redundancy in the information.

Another important point discussed about inbreeding and genomic relationship is how and which SNPs to select to study the relationship between individuals.

Van Raden (2008) reports that it is possible to estimate the relationship of animals based on genetic markers distributed throughout the genome, and shows that a matrix of genomic relationships can be calculated by different methods. Since the genetic similarity between animals can be estimated more accurately based on molecular markers than using pedigree information (Forni, Aguilar & Misztal, 2011). The markers allow to estimate the proportion of loci shared by individuals because it identifies IBD alleles that can be shared by common ancestors not registered in the pedigree (Powell, Visscher & Goddard, 2010). Furthermore, with a high density of markers it has been possible to estimate the proportion of continuous chromosomal segments shared by individuals, identifying IBD haplotypes, through runs of homozygosity (ROH) (Zavarez et al., 2015).

Linkage Disequilibrium

Linkage Disequilibrium (LD) is a non-random association between alleles at different loci, due to physical proximity, population history and evolutionary forces. The extent of LD depends on local recombination rates, for example. In mammals, there are reduced rates of recombination in the sex chromosomes and close to the centromere, while in regions such as euchromatin and small regions known as hotspots this rate of recombination is high (Veroneze, 2015). Places with a high recombination rate cause lower LD and vice versa.

Previous studies indicate that the first step, before running some genomic analyses, is to analyse LD (Santure et al., 2010; Rolf et al., 2010), to exclude linked markers and avoid redundant information. However, there are analyzes where the use of linked markers is considered informative.

Linkage disequilibrium structure can also provide insights into the evolutionary history of a population, breeding system, and pattern of geographic subdivision (Corbin et al., 2010). While the history of natural selection, mutations and other forces that cause changes in genetic frequencies lead to specific differences in LD (Veroneze, 2015).

Effective Population Size

The effective population size (N_e) is described by Falconer & Mackay (1996) as the effective number of individuals that mate and may give rise to the sampling variance or inbreeding rate under a given situation. N_e is an important population parameter related to the amount of genetic variation and genetic drift in a population and represents the minimum number of breeding individuals in an idealized population, showing the same distribution of allele frequencies under random genetic drift or inbreeding as the population under consideration (Barbato et al., 2015).

N_e can be estimated based on pedigree data or genomic data. Regarding genomic data, N_e is estimated taking into account LD patterns (Hayes et al., 2003). Studying the strength of LD at different genetic distances between loci, which allows inferring about an effective ancestral population size that, in turn, can help to understand the influence of breeding strategies on genetic variation present in populations and provide an insight into the level of inbreeding in populations (Corbin et al., 2012; Barbato et al., 2015).

Previous studies show that the extent of LD in animal populations is greater than in humans, mainly due to the small effective population size and typical artificial selection for animal populations (Eusebi, Martinez & Cortes, 2019). LD being at greater recombinant distances is informative about the recent N_e, while shorter distances provide information about ancient N_e (Barbato et al., 2015).

OBJECTIVES

Main Objective

Select a set of SNPs, shared on different commercial SNP chips, to compose an information panel for assigning parentage. In addition to assessing the distribution of inbreeding and genomic relationship coefficients in Gir animals.

Specific Objectives

- **i.** Identify a set of at least 1000 SNP markers, present in the main commercial chips, that accurately allow parentage assignment to be carried out in Gir breed;
- **ii.** Estimate relationship coefficients in Gir cattle based on SNP markers;
- **iii.** Estimate inbreeding coefficients in Gir animals through runs of homozygosity;
- **iv.** Estimate effective population size based on linkage disequilibrium in Gir cattle.

REFERENCES

ARAÚJO, A. M. de. **Parentage and genetic diversity in goats in Brazil using DNA microsatellites**. Doctoral Thesis, Postgraduate Program in Animal Science - Federal University of Viçosa, Viçosa, 2004.

BARBATO, M. et al. SNeP: a tool to estimate trends in recent effective population size trajectories using genome-wide SNP data. **Frontiers in Genetics**, v. 6, p. 109, 2015.

CAETANO, A. R. SNP markers: basic concepts, applications in animal management and breeding and perspectives for the future. **Revista Brasileira de Zootecnia**, vol. 38, n. 8, p. 64-71, 2009.

CORBIN, L. J. et al. Estimation of historical effective population size using linkage disequilibria with marker data. **Journal of Animal Breeding and Genetics**, v. 129, n. 4, p. 257-270, 2012.

CONTI, A.; MORENO, F. S.; ONG, T. P. Nutrigenonics: genomic revolution in nutrition. **Ciência e Cultura**, v. 62, n. 2, p. 04-05, 2010.

EUSEBI, P. G.; MARTINEZ, A.; CORTES, O. Genomic Tools for Effective Conservation of Livestock Breed Diversity. **Diversity**, v. 12, n. 1, p. 8, 2020.

FALCONER, D. S.; MACKAY, T. F. C. **Introduction to Quantitative Genetics**. Essex: Longman Scientific & Technical, 4 ed., 464p, 1996.

FAO. World Food and Agriculture - Statistical, 2021. Available in < http://www.fao.org/faostat/en/#data/QL>. Accessed May 7, 2021.

FLANAGAN, S. P.; JONES, A. G. The future of parentage analysis: From microsatellites to SNPs and beyond. **Molecular Ecology**, v. 28, n. 3, p. 544-567, 2019.

FORNI, S.; AGUILAR, I.; MISZTAL, I.. Different genomic relationship matrices for single-step analysis using phenotypic, pedigree and genomic information. **Genetics Selection Evolution**, v. 43, n. 1, p. 1-7, 2011.

GARANT, D.; KRUUK, L. E. B. How to use molecular marker data to measure evolutionary parameters in wild populations. **Molecular Ecology**, v. 14, n. 7, p. 1843-1859, 2005.

GELDERMANN, H.; PIEPER, U.; WEBER, W. E. Effect of misidentification on the estimation of breeding value and heritability in cattle. **Journal of Animal Science**, v. 63, n. 6, p. 1759-1768, 1986.

HARLIZIUS, B., M. S. LOPES, N. DUIJVESTEIJN, L. H. P. van de GOOR, W. A. van HAERINGEN, H. PANNEMAN, S. E. F. GUIMARÃES, J. W. M. MERKS, E. F. KNOL. A single nucleotide polymorphism set for paternal identification to reduce the costs of trait recording in commercial pig breeding. Journal of Animal Science, Volume 89, Issue 6, P. 1661–1668. 2011.

HAYES, B. J. et al. Novel multilocus measure of linkage disequilibrium to estimate past effective population size. **Genome Research**, v. 13, n. 4, p. 635-643, 2003.

HENDERSON, C. R. Best linear unbiased estimation and prediction under a selection model. **Biometrics**, p. 423-447, 1975.

KALINOWSKI, S. T.; TAPER, M. L.; MARSHALL, T. C. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. **Molecular Ecology**, v. 16, n. 5, p. 1099-1106, 2007.

LOPES, P. S. et al. **Animal Breeding Theory**. Belo Horizonte: FEPMVZ, p. 89, 2005.

MARSHALL, T. C. et al. Statistical confidence for likelihood-based paternity inference in natural populations. **Molecular Ecology**, v. 7, n. 5, p. 639-655, 1998.

NEFF, B. D.; REPKA, J.; GROSS, M. R. A Bayesian framework for parentage analysis: the value of genetic and other biological data. **Theoretical Population Biology**, v. 59, n. 4, p. 315-331, 2001.

NIELSEN, R. et al. Statistical approaches to paternity analysis in natural populations and applications to the North Atlantic humpback whale. **Genetics**, v. 157, n. 4, p. 1673-1682, 2001.

NODARI, R. O. et al. **Biotechnologies**. FIT Handbook 5806 - Biotechnologies. Department of Phytotechnics. Federal University of Santa Catarina. May 2016.

PIERCE, B. A. **Genetics: A Conceptual Approach**. 5th edition. Rio de Janeiro: Guanabara Koogan, 2017.

POWELL, J. E.; VISSCHER, P. M.; GODDARD, M. E. Reconciling the analysis of IBD and IBS in complex trait studies. **Nature Reviews Genetics**, v. 11, n. 11, p. 800-805, 2010.

ROLF, M. M. et al. Impact of reduced marker set estimation of genomic relationship matrices on genomic selection for feed efficiency in Angus cattle. **BMC Genetics**, v. 11, n. 1, p. 24, 2010.

RON, M. et al. Misidentification rate in the Israeli dairy cattle population and its implications for genetic improvement. **Journal of Dairy Science**, v. 79, n. 4, p. 676-681, 1996.

SANTURE, A. W. et al. On the use of large marker panels to estimate inbreeding and relatedness: empirical and simulation studies of a pedigreed zebra finch population typed at 771 SNPs. **Molecular Ecology**, v. 19, n. 7, p. 1439-1451, 2010.

SENNEKE, S. L.; MACNEIL, M. D.; VAN VLECK, L. D. Effects of sire misidentification on estimates of genetic parameters for birth and weaning weights in Hereford cattle. **Journal of Animal Science**, v. 82, n. 8, p. 2307-2312, 2004.

SIGNOROVITCH, J.; NIELSEN, R. PATRI—paternity inference using genetic data. **Bioinformatics**, v. 18, n. 2, p. 341-342, 2002.

SPEED, D.; BALDING, D. J. Relatedness in the post-genomic era: is it still useful? **Nature Reviews Genetics**, v. 16, n. 1, p. 33-44, 2015.

STAFUZZA, N. B. et al. Single nucleotide variants and InDels identified from whole-genome re-sequencing of Guzerat, Gyr, Girolando and Holstein cattle breeds. **PLoS One**, v. 12, n. 3, p. e0173954, 2017.

STALLINGS, R. L. et al. Evolution and distribution of $(GT)_n$ repetitive sequences in mammalian genomes. **Genomics**, v. 10, n. 3, p. 807-815, 1991.

TAYLOR, A. C. et al. Relatedness structure detected by microsatellite analysis and attempted pedigree reconstruction in an endangered marsupial, the northern hairy-nosed wombat Lasiorhinus krefftii. **Molecular Ecology**, v. 6, n. 1, p. 9-19, 1997.

VAN RADEN, P. M. et al. Invited review: Reliability of genomic predictions for North American Holstein bulls. **Journal of Dairy Science**, v. 92, n. 1, p. 16-24, 2009.

VAN VLECK, L. D. Index selection for direct and maternal genetic components of economic traits. **Biometrics**, p. 477-483, 1970.

VERONEZE, R. **Linkage disequilibrium and genomic selection in pigs**. Doctoral Thesis. Wageningen University, 2015.

VISSCHER, P. M. et al. Estimation of pedigree errors in the UK dairy population using microsatellite markers and the impact on selection. **Journal of Dairy Science**, v. 85, n. 9, p. 2368-2375, 2002.

WEIR, B. S.; ANDERSON, A. D.; HEPLER, A. B. Genetic relatedness analysis: modern data and new challenges. **Nature Reviews Genetics**, v. 7, n. 10, p. 771-780, 2006.

ZAVAREZ, L. B. et al. Assessment of autozygosity in Nellore cows (Bos indicus) through high-density SNP genotypes. **Frontiers in Genetics**, v. 6, p. 5, 2015.

Chapter 2

PARENTAGE ASSIGNMENT AND GENOMIC RELATIONSHIP USING DIFFERENT SNP MARKER PANELS IN GIR CATTLE

A. O. Garcia[1], J. C. do C. Panetto[2], M. V. G. B. da Silva[2], P. I. Otto[3], S. E. F. Guimarães[1], and M. A. Machado[2]

[1]Department of Animal Science, Universidade Federal de Viçosa, Viçosa, MG, 36570-900, Brazil

[2]EMBRAPA - Dairy Cattle Research Center, Juiz de Fora, MG, 36038-330, Brazil

[3]Department of Animal Science, Universidade Federal de Santa Maria, Santa Maria, RS, 97105-900, Brazil

ABSTRACT

Pedigree errors considerably reduce the potential genetic progress of the population in question. In this context, our aim was to produce an accurate parentage assignment for implementation in Gir cattle using SNP markers present in four different commercial SNP chips, as well as determining the distribution of genomic relationship coefficients between animals of different classes of family relationship. We used the likelihood ratio approach to reconstruct the pedigree of Gir cattle, in order to correct any errors in annotation and to deepen the information contained in the pedigree records. The approach proved to be satisfactory, improving the depth of the pedigree and showing that it is well established in terms of recent information. In the case of the Gir breed, due to the ease and speed of execution and interpretation, the exclusion method would be the most suitable for routine parentage validation tests. The relationship coefficients allowed us to assess the distribution of values for each relationship category of the reconstructed pedigree.

Keywords: non-imputed, pedigree reconstruction, relatedness, sequoia

INTRODUCTION

In livestock, traditional relatedness coefficients point expected genome sharing between individuals based on pedigree. Measurements made on related individuals can be used to estimate the additive component of variance, which can be estimated from the observed co-variances of quantitative traits between individuals of known relatedness, which in turn were needed to predict the gain from breeding programmes for domesticated animal species. More recent approaches use genetic markers, such as Single Nucleotide Polymorphism (SNP) markers, either to estimate relatedness coefficients with great accuracy in an unobserved pedigree or to identify shared genomic regions that are unaffected by recombination (Speed & Balding, 2014). The observed genotypes can be used to distinguish between a set of possible alternative degrees of relationship, or they can be used to estimate an unknown degree of relationship.

Recently, genotyping using high-density SNP chips has become available at lower cost and it is increasingly being used in genetics research and breeding programs, leading to an increasing number of studies using these markers for parentage validation and assignments (Panetto et al., 2017; Huisman, 2017; Grashei, Ødegård & Meuwissen, 2018). There are currently several commercial genotyping panels available for different species, and it is common for animals of the same breed to be genotyped with chips of different densities.

Nowadays, the simplest approach of molecular parentage analysis is the exclusion method, in which there is a list of candidate parents and the genotypes are compared with the candidate offspring. Genotype compatibility is based on Mendel's inheritance laws, where a minimum percentage of incompatibility is acceptable. In incompatibility events, the candidate parent is removed from the set of candidate parents of that offspring. This method is the most used for its easy execution and interpretation (Flanagan & Jones, 2019). On the other hand, the use of exclusion methods to assign parentage may fail if information is limiting and genotyping errors are present. Statistical approaches based on likelihood methods were developed to overcome these limitations, in which the probabilities of attribution of paternity are determined from simulations. The likelihood method uses of heterozygous genotypes, which are ignored by exclusion methods, reducing the amount of loci required. In addition, can be calculated over many individuals jointly, whereas relatedness is typically calculated pairwise (Huisman, 2017). Both likelihood- and exclusion-based approaches usually assume known and homogenous genotype error rates, do not account for variation in genotype call rates and independent loci (Grashei, Ødegård & Meuwissen, 2018).

Our aim with this study was to produce a reproducible, cost effective and accurate parentage assignment for implementation in Gir cattle using SNP markers present on four different commercial SNP chips, in addition to assessing the distribution of the genomic relationship coefficients between animals for different relationship categories according to inferred pedigree.

MATERIAL AND METHODS

Data

The animals used in this study comprise the data from the Brazilian Dairy Gir Breeding Program (PNMGL), coordinated by Embrapa Dairy Cattle (Juiz de Fora, Minas Gerais, Brazil) in cooperation with the Brazilian Association of Zebu Breeders (ABCZ) and the Brazilian Association of Dairy Gir Breeders (ABCGIL).

The genotypes and genealogy records of 16,983 animals born between 1960 and 2020 were considered, in which the distribution of births can be seen in Figure 1. DNA samples have been analyzed using four different SNP chip densities: 597 animals were genotyped using the Illumina Bovine HD BeadChip with 777K (Illumina, San Diego, CA, USA); 11,207 animals using the GeneSeek Genomic Profiler Indicus (GGPi) BeadChip with 35K; 3,653 animals with the GGPi 50K; and 1,526 animals using the GGPi-LD BeadChip with 27K (GeneSeek Inc., Lincoln, NE, USA). The map positions were determined according to the bovine genome assembly ARS-UCD1.2 (GenBank project accession: NKLS00000000.2).

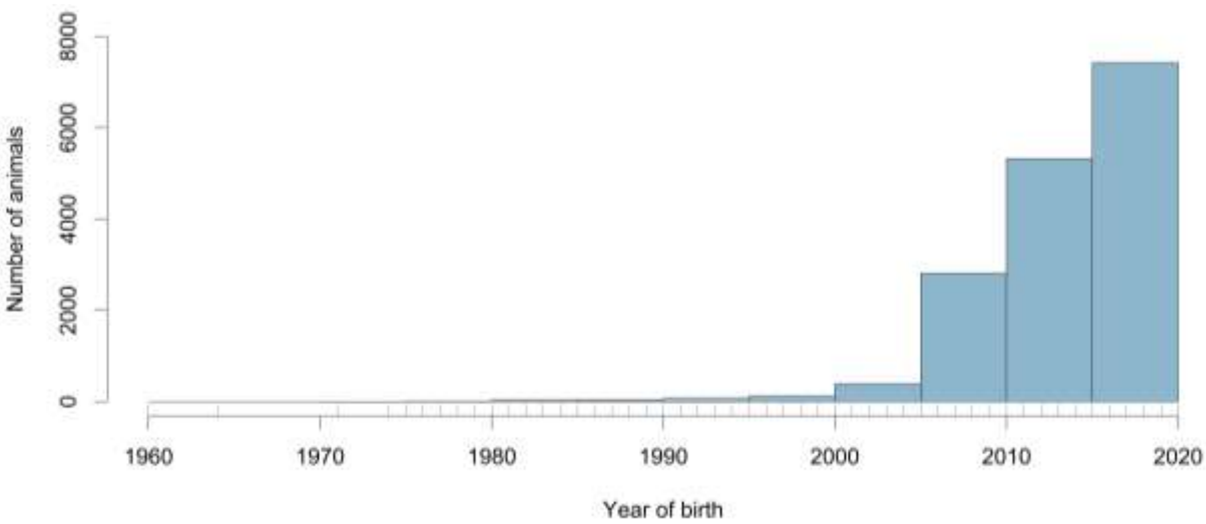


Figure 1. Birth distribution of the 16,983 animals analyzed.

Genotype quality control was implemented using the *snpStats* package (Clayton, 2019) in R software (R version 4.0.2; R Foundation for Statistical Computing, Vienna, Austria). Samples with a call rate < 0.90 and GenCall (GC) score average < 0.70 were removed. Selection of markers across different chips was based on SNP names and their chromosomal positions. SNPs unsigned to any chromosome or mapped to sexual and mitochondrial chromosomes were removed from the dataset. Only autosomal SNPs with call rate > 0.90, deviating from Hardy-Weinberg equilibrium (P value > 10^{-6}), minor allele frequency (MAF) > 0.05 and GC score average > 0.70 present in all four different SNP chips were kept for the later analysis. SNPs were pruned for linkage disequilibrium (LD) to remove linked loci at r^2 threshold > 0.20 using PLINK 1.9 software (Purcell et al., 2007). After quality control, 16,205 Gir animals, 14,458 females and 1,747 males, and 7,372 SNP markers remained. Of these 7,372 SNPs remaining, only the 1,810 common to the four SNP chips were used.

Parentage Analysis and Pedigree Reconstruction

Although we have a pedigree available, for the purposes of this work we consider the worst-case scenario assuming there was no pedigree information; only the genotypes and year of birth of all animals. This entailed that all pairwise relations had to be tested. We used the likelihood method for parentage assignment, described by Marshal et al. (1998). Likelihood-based methods often calculate the likelihood ratio of the genotype of the offspring, which is the probability of the offspring's genotype given the genotypes of the candidate parents, relative to the probability of observing the genotype in the population by chance. Here we use the logarithm base 10 of the likelihood ratios.

Pedigree reconstruction was done using a likelihood-based method, using the *sequoia* package (Huisman, 2017) in R software (R version 4.0.2; R Foundation for Statistical Computing, Vienna, Austria). Likelihood-based methods calculate the likelihood ratio (LR) of the genotype of the offspring, which is the probability of the offspring's genotype given the genotypes of the candidate parents, relative to the probability of observing the genotype in the population by chance. The LR statistic gives more weight to the rare alleles, in addition to considering that all loci are independent, so that the total likelihood ratio is multiplied by all loci (Grashei, Ødegård & Meuwissen, 2018).

A positive log-likelihood ratio (LLR) indicates that this animal being the parent, versus the next most likely relationship between the focal individual and this animal. An LLR value equal to zero implies that the alleged parent is as likely a parent as any animal randomly chosen from the population. While a negative LLR indicates that the hypothesis that individuals are unrelated is more likely, given the genotypes (Marshall et al., 1998; Huisman, 2017; Flanagan & Jones, 2019). Since paternity is attributed to the most likely parent, the output file shows only the relationship between individuals with the highest value of the log-likelihood rate. These LLR values, although useful in obtaining a maximum likelihood solution, cannot be interpreted statistically or biologically.

Assignments were made after considering the likelihoods of all possible first-, second- and third-degree relationships between the focal individuals, as well as the traditional alternative of being unrelated. The following relationships were assigned: Parent-Offspring (PO); Full sibling (FS); Half sibling (HS); Grand-Offspring (GO); Full avuncular (FA); Half avuncular (HA); Full nephew/niece (FN); Half nephew/niece (HN). The distinction between the various categories of second-degree relatives was possible when the probabilities were calculated according to at least one parent of each focal individual.

The quantity that was maximized was the total likelihood ($\mathcal{L}$) of the pedigree configuration over all genotyped individuals (N) (Huisman, 2017):

$$\mathcal{L}(P) = \prod_{A=1}^{N} \mathcal{L}(A, D_A, S_A) \approx \prod_{A=1}^{N} \prod_{l} P(A_l = X | D_A, S_A)$$

where $P(A_l = X | D_A, S_A)$ is the probability of observing genotype X at locus l in individual A, conditional only on its parents D_A and S_A in pedigree P. Which were then multiplied over all individuals and multiplied over all loci.

The *sequoia* package (Huisman, 2017) does not use previous pedigree information, only information about the genotype, sex and year of birth of the individuals. Initially, duplicates, swapped or mislabelled samples were checked, then a series of filtering steps were performed in order to reduce time and computational work: First, the opposite homozygosity occurs, where the number of potential parent–offspring (PO) pairs was drastically reduced; A second filtering step consisted of calculating the log-likelihood ratio (LLR); Finally, the assignments were made based on the highest relationship likelihood and the following criteria: i. an individual cannot be its own ancestor; ii. ancestors are born prior to their descendants, or either or both have an unknown birth year; and iii. the two parents of an individual are of opposite sex, or either one is of unknown sex. The birth years are not used in the LLR calculation and do not affect the value of the likelihoods for the various relationships, but they are used during some filtering steps, and may therefore affect the likelihood ratio.

In the case of domestic populations with a well-established pedigree, *sequoia* package allows an estimate confidence probability through a reference pedigree. Reference pedigree used in this analysis refers to data collected by Brazilian Association of Zebu Breeders (ABCZ) with prior verification of parent-offspring incompatibilities based on the counting of Mendelian conflicts, according to Wiggans et al. (2010), using the SeekParentF90 software (Aguilar, 2014). Threshold number of conflicts to exclude or confirm any parent was set to one percent (1.0%) of the SNP markers. No matching was removed from the reference pedigree. The confidence probability, obtained here through ten simulations, was considered as the number of correct assignments (matching to the reference pedigree), divided by all the assignments made in the simulated pedigrees.

Genomic Relationship Coefficient

The genomic relationship matrix was made available by Embrapa Gado de Leite. In which the imputation was implemented using the FIMPUTE 2.2 software (Sargolzaei, Chesnais & Schenkel, 2014), and lower density panels were imputed to the High Density (HD) level gradually. The animals genotyped with the Bovine HD BeadChip (Illumina Inc., San Diego, CA, USA) were used as reference population for imputation. Imputation accuracy was 0.9669. In this step, 416,155 SNPs were used to build the relationship matrix.

The matrix containing the relationships between all individuals was calculated as standardized correlations, using SNP & Variation Suite Version 8.9.0 (Golden Helix, Bozeman, MT, USA http://www.goldenhelix.com).

RESULTS AND DISCUSSION

Several commercial chips, from different platforms, are being used in the genotyping of farm animals nowadays. The most popular chips in cattle farming are usually those on the Illumina platform, which were initially created based on the European *Bos taurus* genome, which accumulates an ascertainment bias if the chip is used on distantly related breeds (Lachance & Tishkoff, 2013) or in *Bos indicus* animals. Specific chips have been developed for *Bos indicus* cattle, such as the GeneSeek Genomic Profiler Indicus (GGPi) (GeneSeek Inc., Lincoln, NE, USA). The main differences among chips concerns the density of SNPs present in each of them, and the origin of the SNPs, since polymorphic loci are different between taurine and indicine breeds. The GGPi chips were created with the aim of directing and making cheaper the use of chips in indicine breeds and making the technology more accessible to the most diverse users.

In our study, we considered four commercial SNP chips used in Gir cattle animals (Figure 2). The SNPs markers present in all four different chips resulted in a set of 7,372 common SNPs (Figure 2A). After quality control, a panel with 1,810 SNPs common was considered for the parentage testing (Figure 2B).

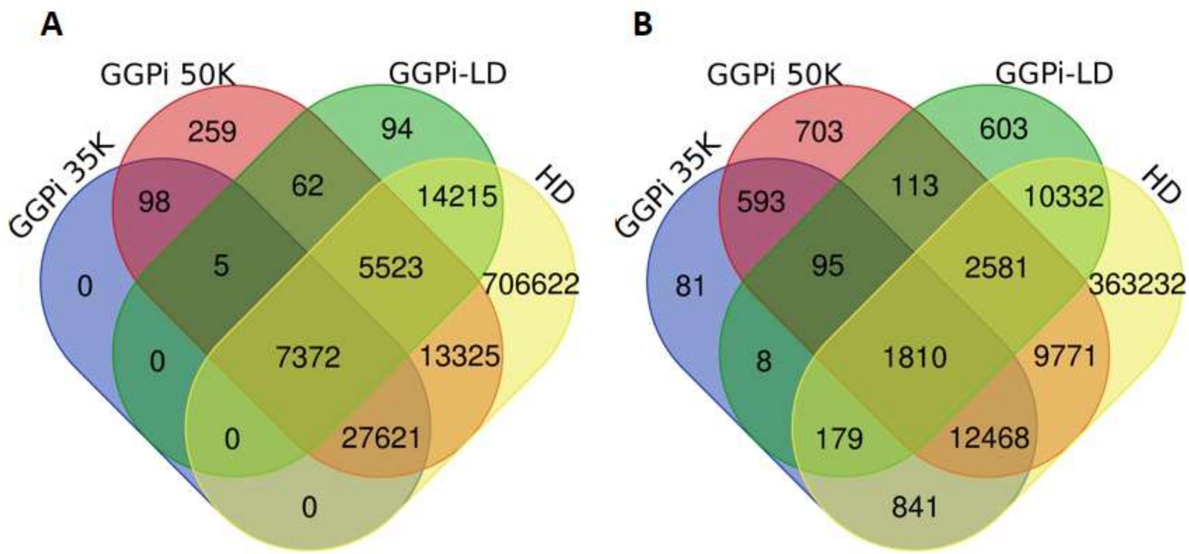


Figure 2. Venn diagram with the SNPs common to the four chips, (A) before quality control and (B) after quality control. Bovine genome sets ARS-UCD1.2 were used. GGPi 35K: GeneSeek Genomic Profiler Indicus (GGPi) with 35K (Geneseek, Lincoln, NE); GGPi 50K: GGPi BeadChip with 50K; GGPi-LD: GGPi-LD BeadChip with 27K; and HD: Illumina Bovine HD with 777K (Illumina, San Diego, CA).

Kopps et al. (2015) conducted a study testing the type and quantity of molecular markers needed to obtain high accuracies in animal relationship tests, concluding that they vary depending on the categories of relationship, mating system and number of overlapping generations. Strucken et al. (2016) found that the minimum number of SNP markers for paternity testing in purebred cattle is at least 200 markers and composite breed cattle is at least 500 markers, with at least 700 SNPs being necessary to avoid attributing incorrect relationships. There are few studies that report parentage testing using high density chips, which is not always feasible to obtain for routine parentage testing due to the cost of genotyping all animals or additional work of imputation, and further investigations on the imputation errors for the test (Wiggans et al., 2009; Calus, Mulder & Bastiaansen, 2011).

The International Society for Animal Genetics (ISAG) recently assembled a SNP panel with 200 markers for international use in parentage tests in cattle. The panel is composed of 100 markers mainly derived from European taurine breeds, and additionally 100 markers selected from indicine and synthetic breeds (ISAG, 2012), therefore, the recommendation is that at least 100 of these markers be used. From the full ISAG panel, 61 markers (3.37% of our panel) were polymorphic in Gir cattle, after quality control. Strucken et al. (2014) in a study with Hanwoo cattle, which despite being also a taurine breed, differed significantly from European taurines, showed that the breed-specific marker panels perform better for the breed than the panel established by ISAG.

Previous studies (Kopps et al., 2015; Dussault & Boulding, 2018; Premachandra, Nguyen & Knibb, 2019) also report the importance of panels with high minor allele frequencies (MAF), values ranging from 0.25 to 0.50, associated with a better chance for correct paternity resolutions, requiring fewer markers. We did not assess the impact of MAFs on parentage assignments, but the values found ranged from 0.06 to 0.50 with average MAF of 0.35 ± 0.12. The descriptive statistics of MAF, as well as the name, chromosome and position of all 1,810 SNPs used in the test, according to the bovine genome assemblies ARS-UCD1.2, are shown in Appendix 1. The distribution of SNPs along the 29 autosome chromosomes can be seen in Figure 3. According to Strucken et al. (2016), the number of SNPs influences more the accuracy of parentage tests than their individual MAF. In addition, a mixture of SNPs with different allele frequencies, resulting in a medium average MAF, seems to work better in parentage tests.

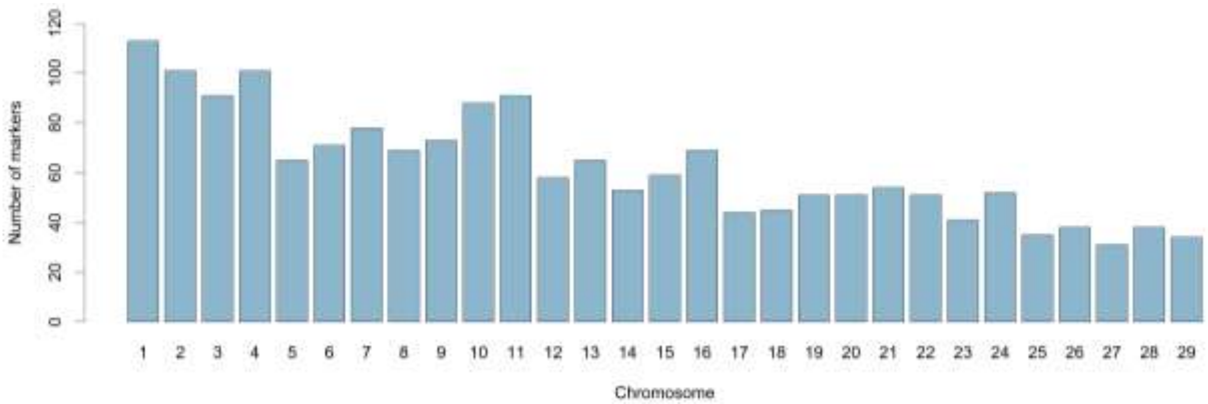


Figure 3. Distribution of 1,810 SPNs common in all four different chips (GeneSeek Genomic Profiler Indicus (GGPi) with 35K and 50K, GGPi-LD BeadChip, and Bovine HD) along 29 autosome chromosomes.

Clarke et al. (2014) emphasize that markers could be selected a priori that show high frequencies in the population, or to apply a test where knowledge about rare alleles can be exploited for better parentage resolution, such as a likelihood ratio.

Parentage Analysis and Pedigree Reconstruction

Analyzing the 16,205 animals and 1,810 SNPs, we found only two negative values of log-likelihood ratio (LLR), being -8.44 for a dam and -1.20 for a sire. The positive values ranged from 3.84 to 84.56 for dams and 1.92 to 128.72 for sires. The average LLR for dams and sires were 42.44 ± 6.08 and 33.77 ± 12.69, respectively. For pairs of parents, values ranged from 36.38 to 89.50, with the average of 63.52 ± 10.77. The distribution of LLR frequencies for pairwise, dams and sires can be seen in Figure 4.

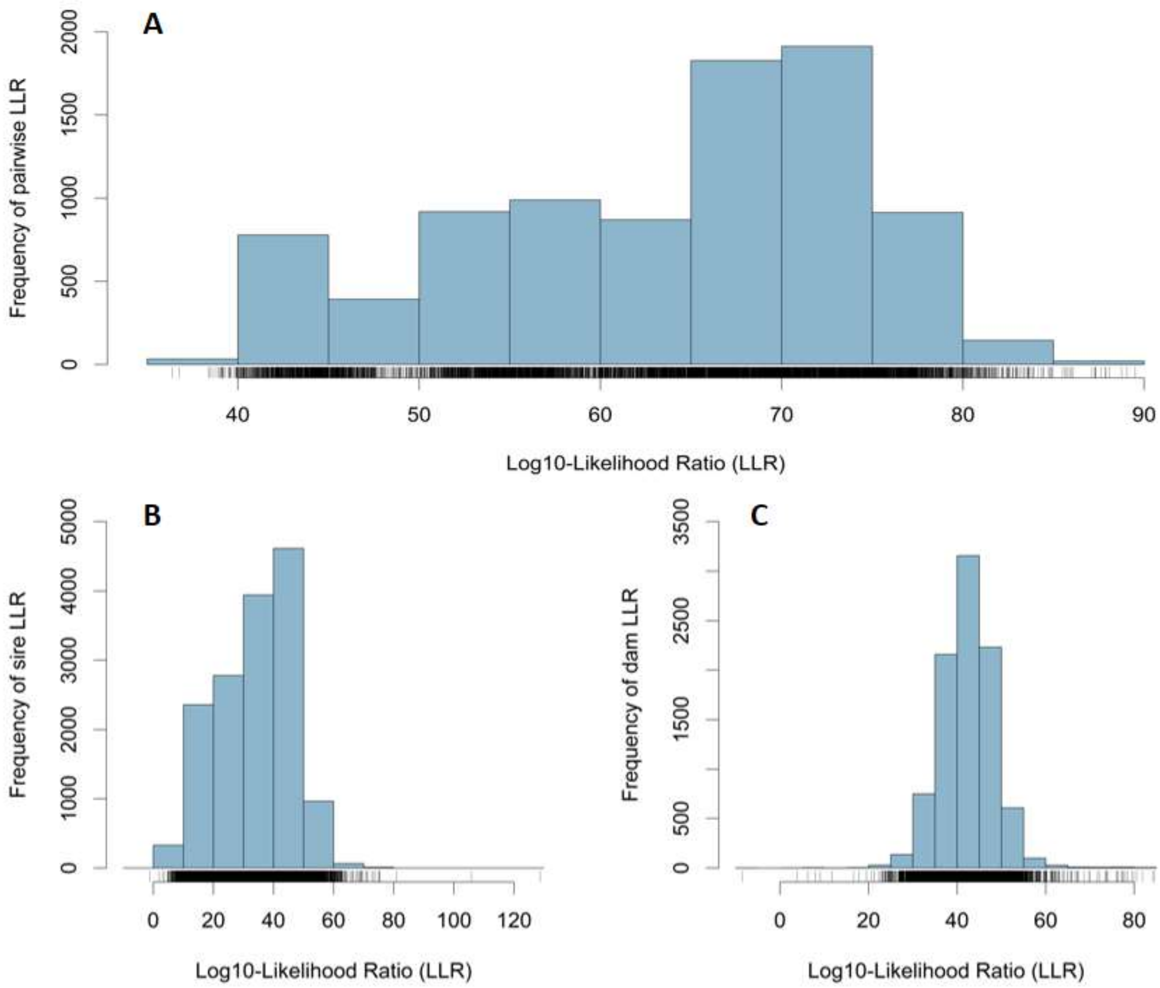


Figure 4. (A) Log10-Likelihood Ratio for the parental pair, versus the next most likely configuration between the three individuals (with one or neither parent assigned). (B) Log10-Likelihood Ratio (LLR) of this male being the sire, versus the next most likely relationship between the focal individual and this male. (C) Idem the previous, for a female parent.

Between the individual and each of the assigned parents the opposite homozygous (OH) locus was checked (Figure 5). In which OH was computed for the parent-offspring relation and Mendelian Error (ME) for trio information (sire-offspring-dam). For ME, the fact that the offspring is OH with both parents is counted as two errors.

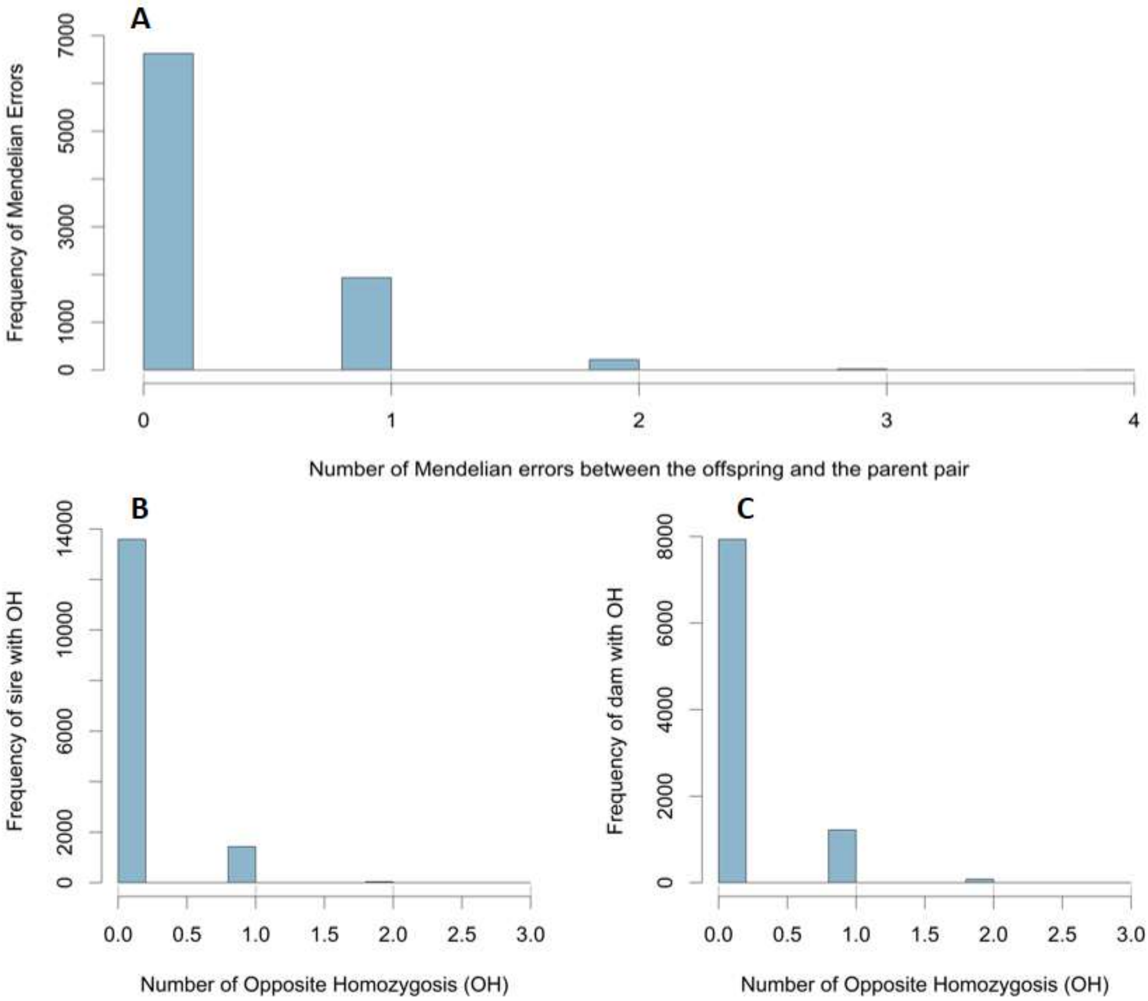


Figure 5. (A) Number of Mendelian errors between the offspring and the parent pair, includes opposite homozygosis, as well as the parents being opposing homozygotes, but the offspring not being a heterozygote. The offspring being opposite homozygous with both parents is counted as two errors. (B and C) Number of loci at which the offspring and sire and dam are opposite homozygotes, respectively.

In true parent–offspring relations, there should be no Mendelian inconsistencies and no locus with OH should occur between the two individuals evaluated, except for mutation events and genotyping errors. Genotyping errors for SNP markers are reported to be generally below 1% (Montgomery et al., 2005; Saunders, Brohede & Hannan, 2007; Wiggans et al., 2009). Of the 9,238 assignments of dams, we observed that 85.92% (n= 7935) did not present opposite homozygosis, 13.21% (n= 1,220) presented OH equal to one and for the sum of OH equal to two and three the percentage was less than 1% (n= 80 and n= 2, respectively). For sires, of the 15,066 paternity assignments, we found 90.22% (n= 13,591) for OH equal to zero, 9.49% for OH equal to one and percentage for the sum of OH equal to two and three represented 0.30% (n= 43 and n= 2, respectively).

For counting ME, only the assignments of pairs of parents were used. The idea is identical to the OH count, in which we count the inconsistencies of only one of the parents, whereas for the ME we consider the inconsistencies of both parents. For pairwise, 8,799 assignments were made, in which 75.33% (n= 6,628) did not present ME, 21.98% (n= 1934) presented one ME, 2.48% presented ME equal to two, and the sum of ME equal to three and four represented only 0.22% (n = 17 and n = 2, respectively).

A simulation analysis was performed to allow the evaluation of the confidence in assignment of parentage to the most likely candidate parent. The confidence probability was taken as the number of assignments considered to be correct, which coincide with the reference pedigree, divided by all the assignments made in the pedigree simulated by the program. Each confidence probability is calculated separately, for sires and dams individually, and for assignments by pair of parents. In this study, we found values of confidence probability equal to 99,9970% for pairs of parents, 99,9980% for sires and 99,9956% for dams.

As previously explained, for the assignment of relationships and reconstruction of the pedigree, some steps prior to calculating the LLR were met, taking into account the date of birth and age differences between animals. The ratio $P(A|R)/P(A)$, which is the ratio between the observed counts of pairs with age difference A and relationship R ($N_{A,R}$), and the expected counts if age and relationship were independent ($N_{.,.} \times p_A \times p_R$). A matrix with the probability ratio of the age difference between two individual's conditionals on them being a certain type of relative ($P(A|R)$) versus being a random draw from the sample ($P(A)$) is generated and is represented by graphical form in Appendix 2. During pedigree reconstruction, the ratios $P(A|R)/P(A)$ calculated were multiplied by the age independent genetic-only

$P(R|G)$ to obtain a probability that the pair are relatives of type R conditional on both their age difference and their genotypes (using Bayes' theorem: $P(A|R)/P(A) \times P(R|G)$).

The comparison between the reference and the inferred pedigree was made using the *PedCompare* command in the *sequoia* package (Huisman, 2017). The summary of the comparison between the pedigrees, as well as the amount and percentage of information contained in them, are shown in Table 1.

Table 1. Comparison between reference (Ref) and inferred (Inf) pedigree, and total information contained in the reference and inferred pedigree for the 16,205 Gir animals.

	Pedigree Compare				**Total Information**			
	Ref_{only}	**Inf_{only}**	**Match**	**Mismatch**	**$Info_{Ref}$**	**$Info_{Inf}$**	**$\%Info_{Ref}$**	**$\%Info_{Inf}$**
Sire-Offspring	6	1,612	13,453	1	13,460	15,066	83.06	92.97
Dam-Offspring	2	1,189	8,049	0	8,051	9,238	49.68	57.01
Sire-Offspring-Dam	3	1,757	7,042	0	7,045	8,799	43.47	54.30

Comparing the reference pedigree and the reconstructed pedigree, we observed that they coincided in 90.01% in the information of sires, 92.65% in the information of dams and 83.98% in trio information (sire-offspring-dams).

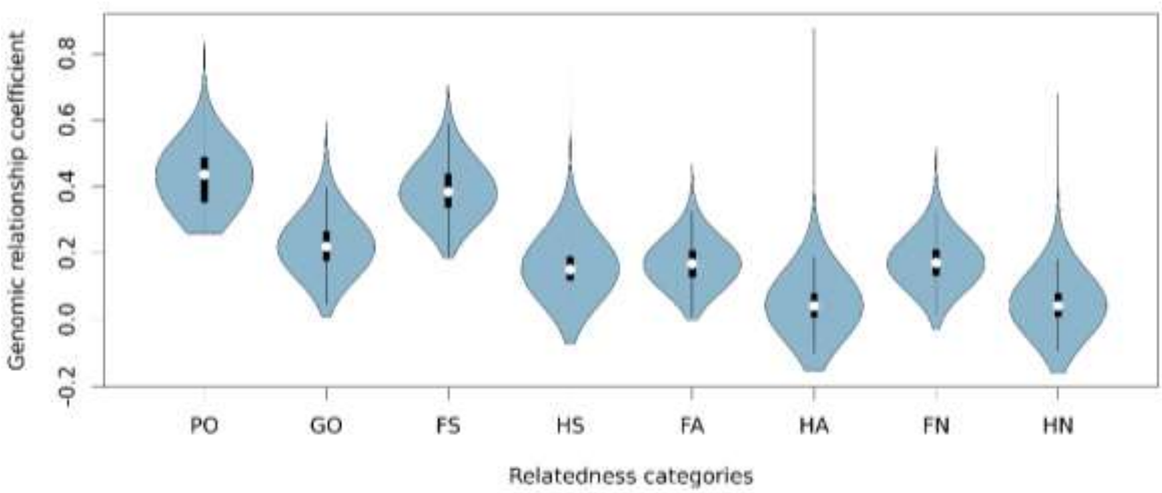
Of the 15,066 sire assignments already mentioned, 1,612 of them occurred only in the inferred pedigree, while in the reference pedigree there were only 13,460 sire information. In the dam category, the inferred pedigree assigned 9,238 dams, while our reference pedigree contained 8,051 dams. In addition, 7.35% of the assignments of dams did not coincide with the reference pedigree. For trio data, the reference pedigree contained 7,045 information, while we managed to assign 8,799 sire-offspring-dams. The trio assignments did not coincide in 16.02% of the cases. The greater number of information in the inferred pedigree may be due to the number of genotyped animals it has, because when there is a large number of genotypes, the likelihood ratio approach forms groups of genotyped relatives (siblings or grandparents) for non-genotyped parents, managing to infer about its genotype.

Even if two individuals are considered as unrelated, according to notes in the pedigree, it is possible to obtain genetic information that supports the hypothesis that they share identical alleles, either by descent or by state. As well as the markers used can detect an incorrect relationship between individuals in a pedigree. The uncertainty is due to the random nature of the parents' allelic inheritance to offspring during the meiosis process and whether the alleles are identical by descent or state, but the use of likelihood ratios allows most of the information about the relationship to be extracted from the observed genotypes. To observe the degree of genome sharing between individuals, considering the pedigree reconstruction performed here, we used the genomic relationship matrix between animals.

Genomic Relationship Coefficient

We prioritize presenting the variation of the genomic relationship coefficient found for each relatedness class.

The *sequoia* R package (Huisman, 2017) allowed us to build a relationship matrix with a dimension equal to NxN (where N= number of animals) from the pedigree file generated in the parentage test carried out previously containing the animal identification, dam and sire using the *GetRelCat* function, where the relationship between individual X and all other individuals in the pedigree is determined going back up to two generations. The following relationships were obtained between all animals: Parent-Offspring (PO); Grand-Offspring (GO); Full Sibling (FS); Half Sibling (HS); Full Avuncular (FA); Half Avuncular (HA); Full Nephew/Niece (FN); and Half Nephew/Niece (HN). From the relatedness class information and the genomic relationship coefficients between the animals it was observed the distribution of these values in each class (Figure 6).



PO= Parent-Offspring; GO= Grand-Offspring; FS= Full Sibling; HS= Half Sibling; FA= Full Avuncular; HA= Half Avuncular; FN= Full Nephew/Niece; and HN= Half Nephew/Niece.

Figure 6. Distribution of the values of coefficient of genomic relationship for each of the eight relatedness categories analyzed.

The descriptive statistics of the genomic relationship values, the number of assignments made by the sequoia software, and the relationship values traditionally expected for each category presented in this work can be seen in Table 2.

Table 2. Descriptive statistics according to the relationship category, with number of assignments ($N_{assignments}$), genomic relationship coefficients (r_{GEN}) and traditional relationship coefficients (r_{PED}) in Gir cattle.

Relatedness categories	**Acronym**	**$N_{assignments}$**	**r_{GEN} min.**	**$\bar{r}_{GEN}$ ± sd***	**r_{GEN} max.**	**r_{PED}**
Parent-Offspring	PO	24304	0.2590	0.4279 ± 0.0826	0.8346	0.50
Grand-Offspring	GO	3096	0.0086	0.2241 ± 0.0691	0.5954	0.25
Full sibling	FS	11900	0.1861	0.3885 ± 0.0726	0.7017	0.50
Half sibling	HS	7691278	-0.0729	0.1564 ± 0.0555	0.7488	0.25
Full avuncular	FA	7969	-0.0022	0.1683 ± 0.0573	0.4654	0.25
Half avuncular	HA	6828787	-0.1526	0.0449 ± 0.0545	0.8764	0.125
Full nephew/niece	FN	9195	-0.0311	0.1732 ± 0.0591	0.5154	0.25
Half nephew/niece	HN	8312857	-0.1624	0.0450 ± 0.0525	0.6788	0.125

*$\bar{r}_{GEN}$ ± sd = average ± standard deviation

As we said in the previous topic, there were 24,304 parentage assignments (dams + sires), with values of the genomic relationship coefficient ranging from 0.259 to 0.835 and average values of 0.428 ± 0.083. There were 3,096 grandparents assignments, for this category the genomic relationship coefficient ranged from 0.009 to 0.595, with an average of 0.224 ± 0.069. For the category of full siblings, there were 11,900 assignments, with a minimum relationship coefficient of 0.1861, an average of 0.388 ± 0.073, and a maximum value equal to 0.7017. While for half siblings, 7,691,278 assignments were made, with relationship coefficient values ranging from -0.073 to 0.749, with an average equal to 0.156 ± 0.056. Unlike the relationship matrix calculated using the pedigree, which contains only positive values, the genomic relationship matrix can contain negative values. This is because we must interpret the values as standardized correlations (VanRaden, 2008; Powell, Visscher & Goddard, 2010). The intuitive explanation is that there is a spread average for these relationships and there are animals that have values above or below this. Animals carrying divergent genotypes will show negative covariances (Legarra, Lourenco & Vitezica, 2018).

The relationship between avuncular and nephew/niece was divided into full and half. The relationship coefficients for the 7,969 full avuncular assignments ranged from -0,002 to 0,465, with an average of 0.168 ± 0.057. While for half avuncular we had 6,828,787 assignments with coefficients ranging from -0,153 to 0,876, and an average of 0,045 ± 0,054. For full nephew/niece 9,195 assignments were made, with a minimum coefficient of -0,031, an average equal to 0,173 ± 0,059, and a maximum value of 0,515. The half nephew/niece category obtained the highest number of assignments, equal to 8,312,857, with average values of genomic relationship coefficient of 0,045 ± 0,052, ranging from -0,162 to 0,679.

CONCLUSION

The panel with 1,810 SNPs proved to be satisfactory for the parentage test, in which the number of assignments by pairwise was equal to 8,799, with confidence equal to 99.9970%. There were 15,065 assignments of sires and 9,237 assignments of dams, with confidence equal to 99.9980% and 99.9956% respectively. Despite the high levels of confidence, the pedigree reconstructed by the likelihood method attributed fewer parents to the individuals than the pedigree that is already used by the Brazilian Association of Zebu Breeders (used as the reference), obtained through manual records made on the farms, indicating that the pedigree used by the association is well established, with reliable and accurate information on the animals' relationship. The methodology added information, which were not in the Association's pedigree, suggesting that when there is a large number of genotyped individuals, the approach is advantageous. Therefore, we recommend the use of this methodology for breeds or species in which there is no well-established pedigree, such as wild species, or breeds in which genealogy is not noted. In the case of the Gir breed, due to the ease and speed of execution and interpretation, the exclusion method would be the most suitable. The relationship coefficients allowed us to assess the distribution of values for each relationship category of the reconstructed pedigree.

ACKNOWLEDGMENTS

We thank EMBRAPA Agricultural Informatics Research Center, the Brazilian Association of Dairy Gir Breeders (ABCGIL), Brazilian Association of Zebu Breeders (ABCZ), and the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Brasilia), and

Ministério da Ciência Tecnologia e Inovação (MCTI)/Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)/Instituto Nacional de Ciência e Tecnologia–Ciência Animal (INCT-CA, Brasilia, Brazil) for financial support for their financial and technical support. The authors have not stated any conflicts of interest.

APPENDIX

Appendix 1. SNP Name, Chromosome (Chr), Position and Minor Allele Frequencies (MAF) of the 1810 SNPs used for parentage testing.

Index	SNP Name	Chr	Position (pb)	MAF
1	BovineHD0100001003	1	3701702	0.4048
2	BovineHD0100001425	1	5060063	0.3678
3	BovineHD0100001461	1	5219539	0.4407
4	BovineHD0100002687	1	9198237	0.1608
5	BovineHD0100002794	1	9559625	0.2777
6	BovineHD0100004065	1	13734238	0.4045
7	BovineHD0100004312	1	14590649	0.4827
8	ARS-BFGL-NGS-22727	1	17041419	0.1529
9	BovineHD0100005671	1	19485236	0.4162
10	BovineHD0100005771	1	19903346	0.3009
11	ARS-BFGL-BAC-11750	1	20285787	0.1458
12	ARS-BFGL-BAC-15578	1	20694788	0.3257
13	BovineHD0100006209	1	21453742	0.4158
14	BovineHD0100006834	1	23629535	0.3815
15	BovineHD0100006953	1	24006741	0.2996
16	BovineHD0100007399	1	25471450	0.4999
17	BTB-00012448	1	29451206	0.4303
18	BovineHD0100008921	1	30743768	0.3186
19	BovineHD0100009460	1	33279806	0.3955
20	BTB-00016647	1	33342088	0.4962
21	BovineHD0100009951	1	35119772	0.3945
22	BovineHD0100010569	1	37344301	0.2966
23	BovineHD0100010664	1	37638872	0.4443
24	BovineHD4100000205	1	37780384	0.3477
25	ARS-BFGL-NGS-86662	1	40576723	0.1392
26	BovineHD0100011903	1	42107409	0.4816
27	BovineHD4100000261	1	47657093	0.3964
28	BovineHD0100014074	1	49639628	0.4709
29	BovineHD0100014223	1	50215571	0.3546
30	BovineHD0100014365	1	50677920	0.1855
31	BovineHD0100014411	1	50857445	0.4993
32	BovineHD0100014467	1	51070788	0.4556

33	BovineHD0100014971	1	52799876	0.2721
34	ARS-BFGL-NGS-18743	1	53214714	0.1867
35	BovineHD0100015665	1	55338266	0.3693
36	BovineHD0100016479	1	57824017	0.2349
37	Hapmap36459-SCAFFOLD9563_462	1	58250592	0.1269
38	BovineHD0100016955	1	59474636	0.4416
39	BovineHD0100017377	1	60783267	0.4474
40	BovineHD0100017508	1	61193765	0.4565
41	BovineHD0100017611	1	61516953	0.4035
42	BovineHD0100017710	1	61908265	0.3793
43	BovineHD0100018065	1	63279200	0.1804
44	ARS-BFGL-NGS-115763	1	65535489	0.1263
45	BovineHD0100018751	1	65896572	0.2554
46	ARS-BFGL-NGS-118306	1	68623214	0.1717
47	BovineHD0100020555	1	71060171	0.3509
48	BovineHD0100021247	1	73499188	0.2766
49	BovineHD0100021467	1	74175764	0.3550
50	BovineHD0100021558	1	74550400	0.2719
51	BovineHD0100021947	1	75664002	0.4669
52	ARS-BFGL-NGS-41839	1	75798317	0.2212
53	ARS-BFGL-NGS-119431	1	78397732	0.0593
54	BTA-38441-no-rs	1	80284430	0.1190
55	BovineHD0100023658	1	81794079	0.4369
56	BovineHD0100024103	1	83588108	0.4225
57	BovineHD0100024179	1	83786080	0.4768
58	BovineHD0100047089	1	86441699	0.2537
59	BovineHD0100025605	1	89587942	0.1592
60	BovineHD0100025708	1	89948634	0.4493
61	BovineHD0100025775	1	90240763	0.3999
62	BTA-94612-no-rs	1	90619077	0.2906
63	BovineHD0100026528	1	93019661	0.4122
64	Hapmap42952-BTA-48143	1	93360980	0.4489
65	Hapmap24383-BTA-158893	1	93876619	0.4889
66	BovineHD4100000546	1	94099296	0.4479
67	BovineHD0100027298	1	95268341	0.4714
68	BovineHD0100027454	1	95845425	0.4505
69	BTB-01469883	1	98550194	0.4967
70	BTA-45203-no-rs	1	99441044	0.2859
71	BovineHD0100028631	1	99491534	0.4783
72	Hapmap39712-BTA-109811	1	100780857	0.1425
73	BovineHD0100029485	1	102605267	0.4970
74	BovineHD0100029576	1	102911722	0.3162
75	BTB-01831301	1	103405790	0.1074
76	BTA-45222-no-rs	1	107500490	0.2003
77	ARS-BFGL-NGS-115502	1	108577099	0.1500
78	BovineHD0100031171	1	109152951	0.4744

79	ARS-BFGL-NGS-105415	1	109939112	0.4586
80	ARS-BFGL-NGS-104897	1	111960854	0.4751
81	BovineHD0100032122	1	112633253	0.4068
82	BovineHD0100032858	1	115313688	0.4490
83	BovineHD0100033499	1	117757291	0.4650
84	BovineHD0100034224	1	119976135	0.3642
85	BovineHD0100034318	1	120350085	0.4046
86	BovineHD0100034485	1	120863450	0.4826
87	BTA-49389-no-rs	1	121510032	0.3424
88	BovineHD0100034908	1	122614173	0.3827
89	BovineHD0100035265	1	124170519	0.4487
90	BovineHD0100035321	1	124287189	0.4586
91	BovineHD0100035695	1	125402936	0.2361
92	ARS-BFGL-NGS-117553	1	130269306	0.4397
93	BovineHD0100047229	1	130344286	0.4443
94	BovineHD0100037698	1	131269582	0.4911
95	BovineHD0100037710	1	131323998	0.3544
96	BovineHD0100037739	1	131451341	0.3889
97	BovineHD0100037763	1	131530420	0.2232
98	BovineHD0100038170	1	133091734	0.4176
99	BovineHD0100039236	1	136350258	0.4453
100	BTA-122684-no-rs	1	137738847	0.1168
101	BTB-00070740	1	138905277	0.4170
102	BovineHD0100040848	1	140784851	0.4622
103	ARS-BFGL-NGS-29228	1	141881684	0.3668
104	BovineHD0100041346	1	142100738	0.4654
105	BovineHD0100041802	1	143223869	0.2177
106	BovineHD0100041883	1	144708430	0.3933
107	BovineHD0100043096	1	147176850	0.3308
108	BovineHD0100043175	1	147357707	0.2334
109	BovineHD0100043192	1	147392045	0.2651
110	BovineHD0100044226	1	150444162	0.3292
111	BovineHD0100044580	1	151406809	0.3541
112	BovineHD0100045028	1	152650093	0.3476
113	BTB-00074258	1	154371107	0.4833
114	Hapmap32829-BTA-147228	2	340299	0.0651
115	ARS-BFGL-NGS-102353	2	3973302	0.1112
116	BovineHD0200001826	2	6291395	0.4082
117	BovineHD0200001880	2	6537154	0.4507
118	BovineHD0200001928	2	6671897	0.2624
119	Hapmap47560-BTA-30470	2	6891780	0.3115
120	BovineHD0200002641	2	9138491	0.3219
121	BovineHD0200002716	2	9556221	0.4818
122	BovineHD0200003044	2	11042964	0.3185
123	BovineHD0200003201	2	11493016	0.4591
124	ARS-BFGL-BAC-35552	2	12802862	0.2470

125	Hapmap59876-rs29018046	2	13822009	0.4016
126	BovineHD0200003972	2	14100829	0.2341
127	BovineHD0200004573	2	15986560	0.4127
128	BovineHD0200005612	2	19470965	0.1993
129	BovineHD0200006991	2	24325865	0.4022
130	BovineHD0200007336	2	25539975	0.4159
131	BovineHD0200007639	2	26277760	0.2823
132	BTB-00091195	2	27364233	0.2848
133	BovineHD0200008343	2	28403750	0.2219
134	BovineHD0200008710	2	29719051	0.3563
135	ARS-BFGL-NGS-64281	2	29852264	0.1691
136	Hapmap51366-BTA-98577	2	30677237	0.4237
137	BovineHD0200009287	2	31265646	0.1624
138	BovineHD0200009736	2	32990258	0.4719
139	BovineHD0200010526	2	35795623	0.4259
140	ARS-BFGL-BAC-30703	2	37052105	0.2038
141	BovineHD0200011184	2	38224098	0.4692
142	BovineHD0200011229	2	38302621	0.4332
143	BovineHD0200011377	2	39047186	0.4676
144	BovineHD0200011483	2	39363095	0.2892
145	BovineHD0200011815	2	40399251	0.3605
146	BovineHD0200012191	2	42042525	0.2847
147	BovineHD0200012239	2	42164025	0.4166
148	BovineHD0200012381	2	42673138	0.3872
149	BovineHD0200012542	2	43344729	0.3821
150	BovineHD0200013899	2	47834929	0.4413
151	BovineHD0200014006	2	48319095	0.4022
152	BovineHD0200040614	2	50955877	0.4691
153	BovineHD0200014747	2	51375334	0.2728
154	BovineHD0200014792	2	51627430	0.2834
155	BovineHD0200014894	2	52048755	0.2431
156	BovineHD0200015162	2	53021140	0.3651
157	BovineHD0200015738	2	55191071	0.4470
158	Hapmap28102-BTA-152636	2	56158911	0.4590
159	BovineHD0200016579	2	58192634	0.1798
160	BovineHD0200016920	2	59243122	0.4353
161	BovineHD0200018446	2	63403696	0.2150
162	BovineHD0200018850	2	64851151	0.4018
163	BovineHD0200019713	2	67873681	0.2934
164	BovineHD0200020536	2	71174131	0.1862
165	BovineHD0200020598	2	71358502	0.4117
166	BovineHD0200020998	2	72871316	0.4805
167	BovineHD0200021341	2	74043645	0.1278
168	BovineHD0200021540	2	74682629	0.4795
169	BovineHD0200021642	2	75032049	0.2282
170	BTB-02066351	2	76142303	0.4855

171	BovineHD0200022609	2	78334369	0.4453
172	BovineHD0200022707	2	78677244	0.4762
173	Hapmap49925-BTA-24427	2	86305851	0.1995
174	BovineHD0200024934	2	87313900	0.1120
175	BovineHD0200025653	2	89893346	0.1579
176	BovineHD0200025721	2	90110992	0.2734
177	ARS-BFGL-NGS-28347	2	90482167	0.3031
178	ARS-BFGL-NGS-33666	2	92377009	0.1979
179	BovineHD0200026726	2	93185827	0.4215
180	BovineHD0200027533	2	95395818	0.1457
181	ARS-BFGL-NGS-115142	2	96938124	0.2178
182	BovineHD0200028080	2	97186892	0.3473
183	Hapmap49348-BTA-85365	2	99409635	0.4933
184	BovineHD0200028910	2	100179964	0.2705
185	ARS-BFGL-NGS-113504	2	100576308	0.3795
186	BovineHD0200029620	2	102678673	0.2913
187	BovineHD0200031021	2	107085658	0.3577
188	BovineHD0200031517	2	108541456	0.1574
189	BovineHD0200031632	2	109025401	0.4893
190	BovineHD0200031897	2	109978168	0.4623
191	BovineHD0200032551	2	112254929	0.4824
192	BovineHD0200032648	2	112527847	0.2598
193	BovineHD0200032760	2	112874857	0.0944
194	BovineHD0200032976	2	113409475	0.3790
195	BovineHD0200033076	2	113657260	0.4432
196	BovineHD0200033092	2	113697220	0.4184
197	ARS-BFGL-NGS-112195	2	116211153	0.3247
198	ARS-BFGL-NGS-67146	2	117962757	0.1859
199	BovineHD0200034945	2	119665548	0.4816
200	BovineHD0200035185	2	120551283	0.2780
201	BovineHD0200035567	2	122156403	0.4670
202	BovineHD0200035909	2	123132894	0.3332
203	BovineHD0200036278	2	124332411	0.2040
204	BovineHD0200036624	2	125526783	0.3267
205	BovineHD0200037020	2	126949310	0.3842
206	BovineHD0200037025	2	126972577	0.0623
207	BovineHD0200037041	2	127012408	0.4096
208	BovineHD0200037069	2	127067457	0.2462
209	BovineHD0200037186	2	127415100	0.3480
210	Hapmap60694-rs29012608	2	128614545	0.0905
211	BovineHD0200037695	2	129095315	0.4179
212	ARS-BFGL-NGS-60038	2	131987746	0.3484
213	ARS-BFGL-NGS-110038	2	133328464	0.2856
214	BovineHD0200039225	2	133834609	0.4212
215	BovineHD0300000045	3	249220	0.4418
216	BovineHD0300000600	3	2117465	0.3943

217	BovineHD0300000602	3	2119650	0.4898
218	BovineHD0300001028	3	3400711	0.4602
219	BovineHD0300035922	3	4452383	0.4734
220	BovineHD0300001719	3	5518576	0.3911
221	BovineHD0300001894	3	6125661	0.4780
222	BovineHD0300002937	3	8789045	0.3485
223	BovineHD0300003770	3	11233396	0.4633
224	ARS-BFGL-NGS-38423	3	12671675	0.3185
225	Hapmap43906-BTA-66658	3	14226434	0.3140
226	BovineHD0300005854	3	18123576	0.4450
227	BovineHD0300006435	3	20262670	0.4444
228	BovineHD0300006597	3	20841126	0.4772
229	BovineHD0300006769	3	21487985	0.3426
230	BovineHD0300007920	3	25032132	0.3717
231	BovineHD0300008058	3	25470563	0.2857
232	BovineHD0300008903	3	27742175	0.3281
233	BovineHD0300009027	3	28172199	0.1190
234	BovineHD0300009414	3	29677682	0.4881
235	ARS-BFGL-NGS-111118	3	31919978	0.4851
236	BovineHD0300010247	3	32719737	0.0919
237	BovineHD0300010258	3	32778065	0.4929
238	BovineHD0300010283	3	32888689	0.4480
239	BovineHD0300010298	3	32909957	0.2692
240	BovineHD0300010316	3	32970665	0.2245
241	Hapmap36453-SCAFFOLD151700_4571	3	33010154	0.0867
242	BovineHD0300010338	3	33060757	0.3096
243	BovineHD0300010725	3	34357923	0.4808
244	BovineHD0300011628	3	37383066	0.2437
245	BovineHD0300012059	3	39269572	0.4071
246	BovineHD0300012355	3	40422105	0.4757
247	BovineHD0300013336	3	43622649	0.2617
248	BovineHD0300013403	3	43814632	0.2872
249	Hapmap43908-BTA-67758	3	44803996	0.1298
250	BovineHD0300013868	3	45335789	0.3290
251	BovineHD0300014097	3	46075566	0.2154
252	BovineHD0300014120	3	46158569	0.4215
253	Hapmap38207-BTA-19427	3	46590052	0.1050
254	BovineHD0300014686	3	48247769	0.4351
255	ARS-BFGL-NGS-114558	3	48432235	0.3304
256	BovineHD0300015009	3	49302622	0.3514
257	ARS-USMARC-Parent-AY842473-rs29001956	3	49548031	0.0802
258	BovineHD0300036114	3	50225937	0.2240
259	Hapmap43654-BTA-67984	3	52242156	0.3140
260	BTA-22647-no-rs	3	52646737	0.2989
261	INRA-417	3	54109833	0.4953

262	Hapmap53284-rs29015774	3	55162621	0.2937
263	BovineHD0300017904	3	59343230	0.4807
264	ARS-BFGL-NGS-119733	3	61330996	0.3930
265	BovineHD0300018468	3	61375577	0.4900
266	BovineHD0300018663	3	62014155	0.4914
267	BovineHD0300018961	3	63135054	0.3732
268	BovineHD0300019502	3	65546461	0.4944
269	Hapmap50441-BTA-68295	3	67294646	0.1746
270	BovineHD0300020358	3	68678422	0.4864
271	Hapmap46316-BTA-68304	3	68825864	0.1494
272	ARS-BFGL-NGS-67327	3	72222667	0.4541
273	ARS-BFGL-NGS-32439	3	75251155	0.4683
274	BovineHD0300021992	3	75492349	0.4168
275	BTB-01508954	3	76404989	0.2048
276	BovineHD0300022303	3	76880284	0.4732
277	BTB-01155362	3	78919364	0.1330
278	BovineHD0300022996	3	79461058	0.4432
279	BovineHD0300023108	3	79981419	0.4036
280	ARS-BFGL-NGS-33471	3	80764030	0.3182
281	BovineHD0300023318	3	80848053	0.3641
282	BovineHD0300035620	3	81854161	0.4899
283	ARS-BFGL-NGS-2215	3	84148683	0.0609
284	BovineHD0300024284	3	84474929	0.4410
285	BovineHD0300024866	3	86303616	0.4035
286	BovineHD0300024982	3	86665690	0.3323
287	BTB-00143272	3	88680107	0.0815
288	BovineHD0300025770	3	89120324	0.4475
289	BovineHD0300026004	3	89803473	0.3636
290	BovineHD0300026151	3	90351343	0.3987
291	BovineHD0300035641	3	91794491	0.4248
292	BovineHD0300028383	3	98156010	0.4494
293	BovineHD0300028610	3	98954216	0.4259
294	ARS-BFGL-NGS-100867	3	99432343	0.2679
295	BovineHD0300029029	3	100845430	0.4854
296	BovineHD0300029520	3	102529569	0.4773
297	BovineHD0300029739	3	103194072	0.2886
298	BovineHD0300030427	3	105525915	0.4237
299	BovineHD0300031492	3	108969717	0.3991
300	BovineHD0300031809	3	109970169	0.4075
301	BovineHD0300032359	3	111814366	0.3876
302	BovineHD0300033420	3	114800677	0.4200
303	BovineHD0300033528	3	115053241	0.4139
304	BovineHD0300033586	3	115240842	0.1719
305	BovineHD0300035025	3	118947879	0.3409
306	BovineHD0400000174	4	1056253	0.4212
307	BovineHD0400000385	4	1760663	0.4891

308	BovineHD0400000439	4	1963960	0.4184
309	BovineHD0400000528	4	2402436	0.4126
310	BovineHD0400001359	4	4929569	0.1896
311	BovineHD0400002768	4	9169569	0.3500
312	BovineHD0400003030	4	10032307	0.2762
313	BovineHD0400003412	4	11406291	0.4591
314	BovineHD0400003748	4	12736935	0.1841
315	ARS-BFGL-NGS-63842	4	13637472	0.3474
316	BovineHD0400005682	4	18986108	0.2713
317	BovineHD0400005756	4	19295757	0.2992
318	BovineHD0400006304	4	20068853	0.4868
319	BovineHD0400006391	4	21481854	0.4715
320	BovineHD0400006759	4	22978697	0.2065
321	BovineHD0400007101	4	24191974	0.3829
322	BovineHD0400035917	4	24647044	0.4296
323	ARS-BFGL-NGS-68962	4	25301015	0.4899
324	BovineHD0400007555	4	25885646	0.4064
325	BovineHD0400007725	4	26679564	0.4903
326	BovineHD0400007769	4	26888454	0.3116
327	BovineHD0400008231	4	28607537	0.3453
328	BovineHD0400008725	4	30671176	0.3654
329	BovineHD0400008815	4	30907393	0.4537
330	BovineHD0400008959	4	31325589	0.4985
331	BovineHD0400009138	4	31922292	0.3733
332	BTA-89083-no-rs	4	33175457	0.1922
333	BovineHD0400009519	4	33449578	0.1937
334	BovineHD0400010171	4	36100661	0.3543
335	BovineHD0400010607	4	37679898	0.4769
336	BovineHD0400010846	4	38485205	0.4906
337	BovineHD0400011142	4	39964119	0.4957
338	BovineHD0400011278	4	40505350	0.4603
339	BTB-01478115	4	41400156	0.0879
340	BovineHD0400011672	4	41800746	0.4160
341	BovineHD0400012014	4	43750599	0.3125
342	BovineHD0400012406	4	45037065	0.3435
343	BovineHD0400012654	4	45941143	0.4719
344	BovineHD0400012684	4	46012758	0.2830
345	BovineHD0400012921	4	46906168	0.2411
346	BovineHD0400036014	4	47335019	0.4689
347	ARS-USMARC-349	4	48001522	0.3242
348	BovineHD0400013452	4	48703644	0.1731
349	BovineHD0400013476	4	48796620	0.3586
350	BovineHD0400013702	4	49426188	0.3900
351	BovineHD0400014626	4	52703446	0.2377
352	BovineHD0400014939	4	53921486	0.4839
353	BovineHD0400015150	4	55182958	0.2250

354	BTB-01889269	4	55367536	0.2822
355	BovineHD0400015641	4	56995593	0.4981
356	BovineHD0400015784	4	57475986	0.4570
357	Hapmap49880-BTA-120276	4	58198819	0.1242
358	ARS-BFGL-NGS-115237	4	60622595	0.3848
359	BTB-00194614	4	66753486	0.4524
360	BovineHD0400018669	4	67485810	0.3138
361	BovineHD0400019532	4	70379038	0.4910
362	Hapmap42648-BTA-71195	4	70610326	0.0663
363	BovineHD0400019722	4	70838504	0.3676
364	BovineHD0400019963	4	71540703	0.2696
365	BovineHD0400020024	4	71792732	0.4882
366	BovineHD0400021096	4	75658211	0.2415
367	ARS-BFGL-NGS-22118	4	76377517	0.4383
368	BovineHD0400021551	4	77148305	0.4366
369	ARS-BFGL-NGS-58613	4	77917876	0.2577
370	BovineHD0400021961	4	78633237	0.4777
371	Hapmap48945-BTA-92136	4	79511525	0.4857
372	BovineHD0400022514	4	80892637	0.2702
373	BovineHD0400023198	4	82963008	0.2689
374	BovineHD0400023288	4	83312875	0.4260
375	BovineHD0400023418	4	83959158	0.4654
376	BTB-01257559	4	85182020	0.1370
377	BTB-01100556	4	86607963	0.2554
378	BovineHD0400024349	4	87025872	0.4397
379	BovineHD0400024919	4	89186231	0.4176
380	BovineHD0400025424	4	90887122	0.4069
381	Hapmap58049-rs29013711	4	91428995	0.3940
382	BovineHD0400026084	4	92519296	0.4672
383	BovineHD0400026486	4	94080994	0.4925
384	BovineHD0400027598	4	97670270	0.4519
385	BovineHD0400027806	4	98440082	0.4843
386	ARS-BFGL-NGS-91439	4	98479733	0.1158
387	BovineHD0400027858	4	98603956	0.4800
388	BovineHD0400028275	4	100064383	0.4502
389	BovineHD0400028562	4	101037390	0.2399
390	Hapmap50076-BTA-71917	4	101104140	0.2809
391	BovineHD0400029079	4	102511490	0.4980
392	BovineHD0400029924	4	105088495	0.4332
393	BovineHD0400030488	4	106366001	0.4985
394	BovineHD0400030825	4	107255314	0.4246
395	BovineHD0400031164	4	108102445	0.4698
396	Hapmap39964-BTA-104515	4	108501357	0.0834
397	Hapmap34657-BES3_Contig465_3267	4	110308532	0.1559
398	BTB-01367046	4	111071819	0.0787
399	BovineHD0400032356	4	111625621	0.4700

400	BovineHD0400032646	4	112540579	0.1521
401	BovineHD0400033124	4	113770744	0.4835
402	BovineHD0400034246	4	116623933	0.4444
403	BovineHD0400034265	4	116662096	0.3247
404	BovineHD0400034787	4	117914528	0.4223
405	BovineHD0400035240	4	119591981	0.1904
406	BovineHD0400035285	4	119766355	0.2022
407	BTB-01735758	5	1104904	0.2597
408	BovineHD0500000702	5	2497981	0.2488
409	BTB-01200617	5	5660281	0.3761
410	BovineHD0500001682	5	6285851	0.4803
411	BovineHD0500001730	5	6397558	0.4713
412	Hapmap22799-BTA-142770	5	6775069	0.3483
413	BovineHD0500002692	5	9436000	0.4753
414	BovineHD0500002788	5	9781802	0.4639
415	BovineHD0500003025	5	10468591	0.4499
416	ARS-BFGL-NGS-117126	5	10802561	0.4449
417	Hapmap32870-BTA-162083	5	14115146	0.2507
418	BovineHD0500004383	5	14778976	0.3621
419	Hapmap44400-BTA-72792	5	15619297	0.0822
420	ARS-BFGL-NGS-20187	5	15943853	0.4163
421	BovineHD0500005433	5	18702083	0.4524
422	BovineHD0500006173	5	21309100	0.4927
423	ARS-BFGL-NGS-72551	5	27324084	0.4929
424	BovineHD0500008750	5	29752408	0.3093
425	BovineHD0500035875	5	31855886	0.0752
426	BovineHD0500009447	5	32387230	0.3301
427	BTB-01205481	5	33629723	0.3692
428	BovineHD0500011229	5	38984821	0.3933
429	Hapmap38177-BTA-73472	5	43622603	0.4129
430	BovineHD0500013381	5	46302236	0.4890
431	BovineHD0500013442	5	46452174	0.1160
432	BovineHD0500015083	5	52297605	0.3843
433	BovineHD0500015178	5	52738353	0.4912
434	BovineHD0500015784	5	55006063	0.1586
435	BovineHD0500017103	5	60746235	0.4819
436	BovineHD0500020197	5	71397484	0.4725
437	BovineHD0500021026	5	73577352	0.3143
438	BovineHD0500021808	5	76020320	0.3752
439	BovineHD0500022029	5	77176583	0.4765
440	BovineHD0500022092	5	77368139	0.4875
441	ARS-BFGL-NGS-21874	5	78895966	0.1927
442	BovineHD0500022723	5	79793030	0.2770
443	BovineHD0500022880	5	80412360	0.2720
444	BovineHD0500022975	5	80748230	0.4660
445	BovineHD0500023327	5	82159800	0.4635

446	BovineHD0500023467	5	82503605	0.4687
447	BovineHD0500023937	5	84177833	0.2817
448	ARS-BFGL-NGS-39913	5	85112695	0.4852
449	BovineHD0500024452	5	85927836	0.4843
450	BovineHD4100003916	5	86523443	0.3548
451	Hapmap48709-BTA-30108	5	86759244	0.2481
452	BovineHD0500024761	5	86942928	0.2534
453	BovineHD0500025127	5	88220440	0.4295
454	BovineHD0500025336	5	88874917	0.4858
455	Hapmap50772-BTA-74551	5	89849731	0.3639
456	BovineHD0500026431	5	92686818	0.4811
457	ARS-BFGL-NGS-116999	5	92796641	0.2523
458	BovineHD0500026762	5	93795810	0.2737
459	BovineHD0500027194	5	95299710	0.2761
460	BovineHD0500027264	5	95634693	0.2472
461	BovineHD0500027388	5	95988609	0.4679
462	BovineHD0500028687	5	99760177	0.1554
463	BovineHD0500035515	5	100887486	0.4210
464	BovineHD0500029252	5	101647536	0.4763
465	BovineHD0500029419	5	102262795	0.4796
466	BovineHD0500030478	5	105724762	0.3354
467	ARS-BFGL-NGS-28247	5	106189372	0.1984
468	ARS-BFGL-NGS-27066	5	107445396	0.2413
469	BovineHD0500031453	5	108671730	0.3210
470	BovineHD0500032263	5	111130674	0.4274
471	BovineHD0500032941	5	113531387	0.4817
472	BovineHD0600000314	6	1301044	0.2807
473	BovineHD0600000793	6	3191815	0.3621
474	ARS-BFGL-NGS-40701	6	6318873	0.4320
475	BovineHD0600001736	6	6590580	0.1699
476	BovineHD0600002048	6	7593285	0.4908
477	BovineHD0600002262	6	8448121	0.0962
478	BovineHD0600002362	6	8809381	0.3058
479	BovineHD0600002514	6	9411599	0.3092
480	BTB-01968603	6	11484420	0.3110
481	Hapmap46194-BTA-103835	6	11714410	0.3267
482	BovineHD0600003873	6	13252803	0.4973
483	BovineHD0600004049	6	13744971	0.3942
484	BovineHD0600005036	6	17162576	0.3177
485	BovineHD0600005319	6	17960875	0.4384
486	BovineHD0600006889	6	23656296	0.4928
487	Hapmap50018-BTA-52623	6	25333843	0.3143
488	BovineHD0600007826	6	26842865	0.3259
489	Hapmap34437-BES1_Contig321_585	6	27154761	0.3000
490	BTA-119035-no-rs	6	30872560	0.3652
491	BovineHD0600009147	6	31233466	0.4801

492	BovineHD0600010035	6	34619116	0.2167
493	BovineHD0600010519	6	36535012	0.4683
494	Hapmap32513-BTC-066089	6	38312868	0.4409
495	BovineHD0600011012	6	38842407	0.3766
496	BTB-02093016	6	42444641	0.2371
497	Hapmap27690-BTC-037987	6	43225858	0.2149
498	BovineHD0600012323	6	43699421	0.4204
499	ARS-USMARC-Parent-DQ789028-rs29017713	6	45364629	0.4979
500	BovineHD0600013064	6	46191554	0.4577
501	ARS-BFGL-NGS-78057	6	47353083	0.2887
502	ARS-BFGL-NGS-58275	6	48335051	0.4092
503	BovineHD0600014453	6	50807764	0.4236
504	BovineHD0600015166	6	53632846	0.4606
505	BovineHD0600015297	6	54217376	0.4691
506	BovineHD0600034716	6	57372560	0.2079
507	BovineHD0600016286	6	57518240	0.3874
508	BovineHD0600016347	6	57825507	0.2370
509	BovineHD0600016410	6	58058134	0.4072
510	BovineHD0600016592	6	58647809	0.4986
511	BovineHD0600017619	6	61987910	0.2972
512	BovineHD0600018460	6	65245056	0.4983
513	BovineHD0600018962	6	67015086	0.4835
514	BovineHD0600019652	6	69104986	0.2104
515	BovineHD0600021024	6	73904776	0.3861
516	BovineHD0600022941	6	81646512	0.3283
517	ARS-BFGL-NGS-114609	6	82992397	0.3175
518	BovineHD0600023757	6	84789192	0.0770
519	BovineHD0600023858	6	85290839	0.1376
520	BovineHD0600023939	6	85748257	0.4869
521	Hapmap38371-BTA-105598	6	85989376	0.4846
522	BovineHD0600024091	6	86371461	0.4055
523	ARS-BFGL-NGS-17376	6	87090740	0.2983
524	BovineHD0600025415	6	90817190	0.4711
525	BovineHD0600025948	6	91926656	0.4065
526	BovineHD0600026278	6	93019191	0.3353
527	BovineHD0600034842	6	95296295	0.4923
528	BTB-00274330	6	96455799	0.1070
529	Hapmap47766-BTA-87827	6	100139940	0.4053
530	BovineHD0600028544	6	100800025	0.2641
531	Hapmap32230-BTC-046313	6	101564868	0.3575
532	BovineHD0600029043	6	102420937	0.2914
533	BovineHD0600029320	6	103121652	0.3701
534	BovineHD0600029617	6	103773464	0.3158
535	Hapmap59570-rs29024349	6	104418701	0.1727
536	BovineHD0600034872	6	106534152	0.4745

537	BovineHD0600031721	6	107226295	0.4529
538	BovineHD0600032184	6	108898352	0.3859
539	BovineHD0600032821	6	110894364	0.3884
540	BovineHD0600032919	6	111164942	0.4658
541	BovineHD0600032991	6	111301803	0.4238
542	ARS-BFGL-NGS-94895	6	116108703	0.2111
543	BovineHD0700033467	7	606235	0.3022
544	BovineHD4100005674	7	1116929	0.4653
545	BovineHD0700000611	7	2378401	0.4395
546	BovineHD0700000656	7	2500957	0.2919
547	BovineHD0700033486	7	4059322	0.4612
548	BovineHD0700001432	7	5216501	0.1256
549	ARS-BFGL-NGS-109521	7	6373887	0.1294
550	ARS-USMARC-Parent-DQ888309-rs29013741	7	7248738	0.2207
551	BovineHD0700002338	7	8099688	0.3675
552	BovineHD0700003027	7	10602763	0.4184
553	ARS-BFGL-NGS-54629	7	13025016	0.1097
554	BovineHD0700004028	7	13740271	0.4138
555	BovineHD0700004129	7	13981000	0.4562
556	BovineHD0700004491	7	15014310	0.1546
557	ARS-BFGL-NGS-118411	7	16406408	0.2588
558	BovineHD0700005111	7	16976604	0.2828
559	BovineHD0700005515	7	18569768	0.4995
560	BovineHD0700005782	7	19601934	0.3758
561	BovineHD0700005847	7	19889676	0.4241
562	BovineHD0700006134	7	21006797	0.3025
563	ARS-BFGL-NGS-250	7	24706140	0.4399
564	BTA-78663-no-rs	7	25318035	0.1640
565	BovineHD0700007727	7	26403007	0.0987
566	BovineHD0700007789	7	26518740	0.2852
567	ARS-BFGL-NGS-11310	7	27072148	0.1438
568	ARS-BFGL-NGS-29250	7	28049997	0.2500
569	BTA-26849-no-rs	7	28462618	0.2307
570	BovineHD0700008485	7	28787927	0.2666
571	BovineHD0700008719	7	29491899	0.3140
572	BovineHD0700009213	7	30971175	0.3996
573	BovineHD0700009378	7	31429263	0.4296
574	BTB-00304608	7	36626525	0.2225
575	BovineHD4100005953	7	37707932	0.4804
576	BovineHD0700011395	7	37957194	0.2292
577	BovineHD0700011845	7	39517939	0.4024
578	BovineHD0700011903	7	39743104	0.3796
579	BovineHD0700012333	7	40941529	0.3963
580	BovineHD0700013934	7	46618431	0.3095
581	BovineHD0700014512	7	48503611	0.4096

582	BovineHD0700014550	7	48588660	0.4450
583	BovineHD0700015341	7	51678528	0.2243
584	BovineHD0700016515	7	55795772	0.2909
585	BTB-00316348	7	62936236	0.3064
586	ARS-BFGL-NGS-116814	7	65083069	0.4238
587	BTB-01668289	7	66626152	0.2273
588	ARS-BFGL-NGS-75486	7	67001032	0.4247
589	BovineHD0700020727	7	68447490	0.4115
590	BovineHD0700021141	7	69682207	0.3069
591	ARS-BFGL-NGS-107513	7	72828445	0.2877
592	BovineHD0700022541	7	74699455	0.2306
593	Hapmap26748-BTA-15768	7	75473623	0.1433
594	ARS-BFGL-NGS-97447	7	75677689	0.4867
595	BovineHD0700022881	7	75905358	0.3207
596	BTA-23130-no-rs	7	76464440	0.1895
597	BovineHD0700023971	7	79883158	0.4249
598	BovineHD0700024538	7	81326759	0.3331
599	ARS-BFGL-NGS-2769	7	81343677	0.0782
600	BTB-01141010	7	84051130	0.3442
601	BovineHD0700025386	7	84242576	0.4626
602	BovineHD0700026301	7	87353780	0.2894
603	BovineHD0700027119	7	90433821	0.3279
604	ARS-BFGL-NGS-100632	7	93176056	0.3329
605	BovineHD0700028016	7	93767005	0.3888
606	Hapmap23231-BTA-80214	7	93916381	0.0841
607	BovineHD0700028121	7	94146131	0.4873
608	BovineHD0700029984	7	100226847	0.4128
609	Hapmap31441-BTA-145417	7	100239773	0.4238
610	BovineHD0700030132	7	100808221	0.4726
611	BovineHD0700030519	7	102072791	0.4871
612	BovineHD0700030586	7	102317457	0.3480
613	BTB-00328499	7	104068858	0.3757
614	BovineHD0700031119	7	104089298	0.4593
615	BTB-00330149	7	106742506	0.4373
616	BovineHD0700031900	7	106818911	0.3339
617	BovineHD0700034068	7	107921604	0.4354
618	ARS-BFGL-NGS-55124	7	109160716	0.3414
619	BovineHD0700032839	7	109810616	0.3304
620	BovineHD0700033246	7	110011927	0.2409
621	BovineHD0800000061	8	460032	0.4092
622	Hapmap51143-BTA-111274	8	1929744	0.2734
623	BovineHD0800002111	8	6736640	0.3281
624	BovineHD0800002962	8	9366638	0.1831
625	BovineHD0800003495	8	10862272	0.3643
626	BovineHD0800004477	8	14170708	0.4753
627	Hapmap32663-BTA-30828	8	14820157	0.4502

628	BTB-00336457	8	16435624	0.1837
629	BovineHD0800006014	8	19395945	0.3493
630	BovineHD0800033609	8	21280749	0.4562
631	BTB-01896501	8	22847354	0.1575
632	BovineHD0800007182	8	23999018	0.4280
633	BovineHD0800007267	8	24229478	0.3160
634	BovineHD0800033638	8	25394398	0.2928
635	Hapmap33212-BTA-145669	8	31048572	0.2395
636	BovineHD0800009695	8	32045406	0.3255
637	BTB-01356324	8	34728021	0.4404
638	BovineHD0800010648	8	35746007	0.3318
639	BovineHD0800011396	8	37966536	0.3989
640	BovineHD0800011629	8	38864815	0.4875
641	BovineHD0800011695	8	39184200	0.3849
642	BovineHD0800012611	8	41994960	0.4285
643	BovineHD0800012680	8	42360502	0.4716
644	BovineHD0800013579	8	45243954	0.3925
645	BovineHD0800014238	8	47235746	0.4662
646	BovineHD0800014403	8	47700446	0.3809
647	Hapmap41647-BTA-81135	8	49711360	0.2926
648	BovineHD0800015415	8	51140820	0.4647
649	BovineHD0800017133	8	56696076	0.4460
650	BovineHD0800017699	8	58823786	0.4952
651	BovineHD0800017756	8	59026861	0.4226
652	BovineHD0800017914	8	59650629	0.3178
653	ARS-BFGL-NGS-115798	8	60447874	0.4867
654	BovineHD0800019134	8	63465667	0.4653
655	BTB-00950285	8	65148173	0.1791
656	BovineHD0800020247	8	67007571	0.4057
657	ARS-BFGL-NGS-11076	8	67672508	0.2248
658	BovineHD0800020777	8	68639303	0.4748
659	BovineHD0800021867	8	71848112	0.2024
660	BovineHD0800021940	8	72132617	0.4045
661	BTB-00357728	8	72833245	0.2247
662	BovineHD0800022997	8	75407498	0.3669
663	UA-IFASA-2457	8	77387217	0.1990
664	BTA-90142-no-rs	8	78433277	0.1560
665	BovineHD0800024036	8	79124368	0.4959
666	BovineHD0800024417	8	80784254	0.4609
667	ARS-BFGL-NGS-108549	8	80923031	0.4989
668	BTB-01092452	8	81078533	0.3717
669	BovineHD0800024653	8	81419378	0.3215
670	BovineHD0800025154	8	83153169	0.2668
671	BovineHD0800025446	8	84381147	0.4876
672	BovineHD0800025895	8	85890528	0.2826
673	BovineHD0800025964	8	86040652	0.4633

674	BovineHD0800026383	8	87457398	0.1484
675	BovineHD0800026510	8	87878438	0.4730
676	BovineHD0800027115	8	89710764	0.4485
677	BovineHD0800027661	8	91496290	0.4856
678	BovineHD0800027865	8	92293700	0.4619
679	BovineHD0800028075	8	93241971	0.4189
680	BovineHD0800028644	8	95256054	0.4909
681	BovineHD0800028825	8	95901419	0.2016
682	ARS-BFGL-NGS-34028	8	99197740	0.4754
683	BovineHD0800030195	8	99979399	0.2700
684	Hapmap55441-rs29010990	8	101431379	0.0698
685	ARS-BFGL-NGS-115630	8	103261302	0.1125
686	BovineHD0800032580	8	106680732	0.2969
687	BovineHD0800032739	8	107163526	0.4128
688	ARS-BFGL-NGS-92824	8	109827619	0.2134
689	ARS-BFGL-NGS-111473	8	110915706	0.3572
690	BovineHD0900030996	9	395020	0.4611
691	BovineHD0900000871	9	4382323	0.4430
692	BovineHD0900001235	9	5826144	0.1704
693	BovineHD0900002667	9	10478676	0.0676
694	BovineHD0900002670	9	10497912	0.3425
695	BovineHD0900002672	9	10501899	0.2533
696	BovineHD0900002673	9	10503858	0.1744
697	BovineHD0900002743	9	10794461	0.4919
698	ARS-BFGL-NGS-59162	9	12492323	0.4309
699	BovineHD0900003395	9	12785651	0.4582
700	BovineHD0900003551	9	13362606	0.4545
701	ARS-BFGL-NGS-14740	9	14224642	0.4928
702	BovineHD0900003794	9	14367327	0.2930
703	BovineHD0900003919	9	14784037	0.4847
704	BTB-00380899	9	15846479	0.3031
705	BovineHD0900004322	9	15946986	0.4665
706	BovineHD0900004661	9	16970210	0.4327
707	BovineHD0900004794	9	17481596	0.3070
708	BovineHD0900004954	9	18054205	0.4961
709	BTB-01964666	9	20780517	0.1017
710	BovineHD0900005768	9	21183118	0.4439
711	BovineHD0900007218	9	26452726	0.4209
712	ARS-BFGL-NGS-93119	9	27778476	0.2736
713	BovineHD0900007855	9	28598682	0.4457
714	BovineHD0900008044	9	29323250	0.4683
715	BovineHD0900009322	9	33496695	0.4880
716	BovineHD0900009478	9	34158326	0.4760
717	BTB-00124923	9	34935085	0.0585
718	ARS-BFGL-NGS-22883	9	36491907	0.3596
719	BovineHD0900010404	9	36933151	0.4278

720	BovineHD0900010596	9	37690979	0.3232
721	Hapmap29482-BTA-146449	9	38324504	0.0964
722	Hapmap42860-BTA-17231	9	39917723	0.4191
723	BovineHD0900012467	9	44373784	0.1844
724	Hapmap24524-BTA-107865	9	45754622	0.1223
725	ARS-BFGL-NGS-31640	9	51702071	0.1056
726	BTB-00392898	9	53249636	0.4465
727	BovineHD0900015381	9	55234581	0.4678
728	BTB-01152124	9	56541892	0.0933
729	Hapmap50486-BTA-83799	9	57663523	0.1825
730	BovineHD0900016605	9	59617973	0.4914
731	BTB-00397198	9	60717471	0.4949
732	BovineHD0900017659	9	63433228	0.4911
733	BovineHD0900018476	9	66009549	0.2565
734	BovineHD0900018538	9	66144440	0.4740
735	BovineHD0900019166	9	68381769	0.4360
736	ARS-BFGL-NGS-31553	9	71565043	0.2610
737	BovineHD0900020630	9	73185420	0.4699
738	BovineHD0900020781	9	73709388	0.4624
739	BovineHD0900020901	9	74171669	0.1666
740	Hapmap33532-BTA-84282	9	74590074	0.3447
741	BovineHD0900021133	9	74925057	0.4386
742	BovineHD0900021957	9	77835234	0.1324
743	BovineHD0900022517	9	80057313	0.2901
744	BTA-84564-no-rs	9	84464824	0.0681
745	BovineHD0900031601	9	86580735	0.4275
746	BovineHD0900025077	9	87793968	0.4706
747	BTB-00404639	9	88877587	0.2722
748	BovineHD0900025958	9	90593399	0.4036
749	BovineHD0900026421	9	91988897	0.3265
750	BovineHD0900026463	9	92152252	0.4714
751	BovineHD0900026546	9	92378566	0.2788
752	ARS-BFGL-NGS-55179	9	94022498	0.1429
753	BovineHD0900027233	9	94324071	0.4883
754	ARS-BFGL-NGS-113972	9	94740322	0.2487
755	BovineHD0900027627	9	95484291	0.2912
756	BovineHD0900028097	9	96651231	0.2985
757	BovineHD0900028738	9	97923106	0.4953
758	ARS-BFGL-NGS-5845	9	98519942	0.3436
759	BovineHD0900029449	9	99744276	0.4815
760	ARS-BFGL-NGS-22624	9	101032356	0.4499
761	BovineHD0900030186	9	101581038	0.4466
762	BovineHD0900030604	9	102689188	0.4973
763	BovineHD1000000349	10	1237165	0.3108
764	Hapmap55445-rs29014484	10	2032397	0.2676
765	BovineHD1000000632	10	2159228	0.4941

766	ARS-BFGL-NGS-33007	10	5602819	0.1702
767	BovineHD1000002662	10	8130277	0.3866
768	BovineHD1000002966	10	9104228	0.4868
769	BovineHD1000003115	10	9499334	0.4520
770	BovineHD4100007862	10	11817116	0.3309
771	ARS-BFGL-NGS-92296	10	12465372	0.4717
772	BovineHD1000004631	10	13818636	0.4715
773	ARS-BFGL-NGS-11874	10	14415124	0.4706
774	BovineHD1000004917	10	14832924	0.3233
775	ARS-BFGL-NGS-119197	10	15189729	0.4404
776	BovineHD1000005587	10	16751413	0.4655
777	BovineHD1000005671	10	16962501	0.4032
778	BovineHD1000005844	10	17441684	0.4097
779	ARS-BFGL-NGS-22915	10	19605083	0.0907
780	BovineHD1000006923	10	21409146	0.3980
781	BovineHD1000008961	10	27248755	0.4076
782	BovineHD1000009028	10	27488970	0.4820
783	Hapmap48539-BTA-97781	10	28936022	0.4775
784	BovineHD1000009834	10	29737800	0.3896
785	Hapmap32413-BTA-63331	10	29994675	0.4889
786	BTB-00416964	10	31244060	0.1924
787	BovineHD1000010217	10	31598764	0.4523
788	BovineHD1000010316	10	32159525	0.4747
789	BovineHD1000010339	10	32238126	0.4223
790	BovineHD1000010586	10	33134088	0.3161
791	BovineHD1000010861	10	34271738	0.4771
792	BovineHD1000011175	10	35427508	0.2815
793	BovineHD1000011755	10	37910321	0.1623
794	BovineHD1000012025	10	38856531	0.4655
795	BovineHD1000012209	10	39592880	0.4156
796	ARS-BFGL-NGS-109834	10	40717249	0.1611
797	BovineHD1000012603	10	41106719	0.4042
798	Hapmap34990-BES11_Contig306_895	10	41919935	0.4522
799	BovineHD1000012963	10	42883250	0.4791
800	BovineHD1000013553	10	44961802	0.2887
801	BovineHD1000014025	10	46724768	0.4398
802	BovineHD1000014406	10	47839131	0.4744
803	ARS-BFGL-NGS-29968	10	48249714	0.1079
804	BovineHD1000014946	10	49507737	0.2818
805	BovineHD1000015174	10	50351945	0.4100
806	BovineHD1000015647	10	52334482	0.4696
807	BovineHD1000015861	10	52989208	0.4497
808	BTB-00430410	10	55670655	0.1526
809	BovineHD1000016652	10	55938538	0.4508
810	BovineHD1000016709	10	56176597	0.4503
811	Hapmap53237-rs29021167	10	62870803	0.4250

812	BovineHD1000019236	10	66675013	0.4691
813	BovineHD1000019314	10	67089148	0.2098
814	BovineHD1000030727	10	67189517	0.2914
815	BovineHD1000019446	10	67625813	0.4455
816	BovineHD1000019518	10	67906788	0.4928
817	BovineHD1000019839	10	68856993	0.3502
818	BovineHD1000020490	10	71272989	0.4617
819	BovineHD1000020947	10	73476783	0.4423
820	BovineHD1000020980	10	73524530	0.3943
821	BovineHD1000021608	10	75587492	0.3229
822	BovineHD1000022732	10	79290718	0.2922
823	BTB-00437636	10	80612799	0.3082
824	BovineHD1000030784	10	81311061	0.2721
825	BovineHD1000023728	10	82931835	0.4901
826	BovineHD1000030805	10	85182618	0.2303
827	BovineHD1000024601	10	85908659	0.2808
828	BovineHD1000024654	10	86141080	0.4747
829	BovineHD1000025484	10	88472859	0.4504
830	BovineHD1000025652	10	88915793	0.4788
831	BovineHD1000025741	10	89062089	0.4470
832	BovineHD1000025838	10	89298637	0.2982
833	BovineHD1000026069	10	89879589	0.4068
834	BTB-00444654	10	90648336	0.1179
835	BovineHD1000026452	10	90669030	0.4584
836	Hapmap47899-BTA-26090	10	91769676	0.4224
837	BovineHD1000027113	10	92750257	0.2410
838	ARS-BFGL-NGS-67195	10	94377386	0.4478
839	BovineHD1000027635	10	94588342	0.3407
840	ARS-BFGL-NGS-68683	10	95898502	0.1577
841	BovineHD1000028127	10	96253449	0.4935
842	BovineHD1000028263	10	96712990	0.3914
843	BovineHD1000028343	10	97065344	0.3385
844	Hapmap39952-BTA-86345	10	98561555	0.1780
845	Hapmap54967-rs29014597	10	99928192	0.2153
846	BTB-01535139	10	100386078	0.3984
847	BovineHD1000030214	10	102348699	0.3375
848	BovineHD1000031631	10	102773611	0.3667
849	BovineHD1000030428	10	103009760	0.2376
850	BovineHD1000030474	10	103103253	0.2548
851	BovineHD1100000047	11	230785	0.3041
852	BovineHD1100000664	11	1934519	0.3715
853	BovineHD1100000942	11	2712184	0.4101
854	BovineHD1100001301	11	3571192	0.4763
855	BovineHD1100002079	11	5753058	0.3201
856	BovineHD4100008518	11	6696865	0.2367
857	ARS-BFGL-NGS-35216	11	7595786	0.4341

858	BovineHD1100002887	11	7787803	0.1857
859	BovineHD1100002939	11	8005521	0.3161
860	BovineHD1100003564	11	10381486	0.4841
861	ARS-BFGL-NGS-22092	11	11412114	0.2615
862	BovineHD1100004154	11	12686068	0.4013
863	BovineHD1100004709	11	14514622	0.1968
864	BovineHD1100004908	11	15329636	0.3244
865	BovineHD1100005014	11	15722386	0.3184
866	ARS-BFGL-NGS-21563	11	16384871	0.2824
867	BovineHD1100005273	11	16902428	0.4886
868	BTB-01391227	11	17038154	0.3270
869	BovineHD1100006223	11	20619336	0.4615
870	BovineHD1100006270	11	20783706	0.4949
871	ARS-BFGL-NGS-26022	11	21626518	0.3634
872	BovineHD1100006654	11	22050528	0.4577
873	BTB-00466422	11	22900356	0.1576
874	BovineHD1100007026	11	23254049	0.4064
875	BovineHD1100007430	11	24646345	0.4296
876	BovineHD1100008456	11	28610612	0.2428
877	BovineHD1100009184	11	30915406	0.3297
878	BovineHD1100009423	11	31476812	0.3346
879	BovineHD1100010325	11	34999550	0.3523
880	BovineHD1100010807	11	36697881	0.2918
881	ARS-BFGL-NGS-66512	11	37131360	0.4143
882	BovineHD1100011314	11	38450680	0.2699
883	Hapmap25293-BTA-148337	11	39555856	0.4417
884	BovineHD1100012377	11	42434080	0.3911
885	BovineHD1100013031	11	45035956	0.2616
886	ARS-BFGL-NGS-112202	11	45290515	0.4573
887	BovineHD1100014206	11	48537236	0.4355
888	BovineHD1100014577	11	49789223	0.2268
889	BovineHD1100014719	11	50235870	0.2247
890	BTB-00474688	11	50529365	0.2922
891	BovineHD1100014985	11	51003233	0.3809
892	BovineHD1100015113	11	51502570	0.3116
893	Hapmap32633-BTA-157234	11	53855205	0.1912
894	BovineHD1100015690	11	53917743	0.4674
895	BovineHD1100015955	11	54983163	0.4471
896	BovineHD1100017310	11	60846941	0.4468
897	BovineHD1100031287	11	61921507	0.2194
898	BovineHD1100017649	11	62320220	0.3478
899	BovineHD1100017708	11	62565516	0.4729
900	BovineHD1100017798	11	62796689	0.2930
901	ARS-BFGL-NGS-119320	11	64307937	0.1590
902	BovineHD1100018471	11	65443157	0.1850
903	BovineHD1100018515	11	65598144	0.4757

904	BovineHD1100018725	11	66224160	0.4647
905	BovineHD1100018784	11	66465014	0.4163
906	BovineHD1100018962	11	67181433	0.2505
907	BovineHD1100019084	11	67551035	0.4938
908	BovineHD1100019212	11	67949488	0.3862
909	BovineHD1100019382	11	68637750	0.3590
910	BovineHD1100020235	11	70610606	0.3378
911	BovineHD1100021092	11	73825531	0.4848
912	BovineHD1100021901	11	76492079	0.4319
913	BovineHD1100021992	11	76779895	0.4598
914	Hapmap51377-BTA-106302	11	76863768	0.1545
915	BovineHD1100031314	11	78846668	0.3572
916	BovineHD1100022662	11	78994191	0.4770
917	Hapmap27739-BTA-126803	11	79945955	0.3183
918	BovineHD1100023014	11	80274899	0.4610
919	BovineHD1100023185	11	80764101	0.3400
920	BovineHD1100023507	11	81906564	0.4940
921	BovineHD1100023541	11	82001946	0.1805
922	BovineHD1100023576	11	82152417	0.4733
923	ARS-BFGL-NGS-5662	11	83049429	0.2183
924	BTA-108440-no-rs	11	84097016	0.2262
925	BovineHD1100024849	11	86517245	0.3475
926	BovineHD1100024917	11	86850436	0.3948
927	ARS-BFGL-NGS-83535	11	89047690	0.3312
928	BovineHD1100025772	11	89368698	0.2329
929	BovineHD1100026702	11	91977613	0.4547
930	BTA-17666-no-rs	11	92079563	0.3516
931	BovineHD1100027156	11	93232642	0.3735
932	ARS-BFGL-NGS-1846	11	93606537	0.3075
933	BovineHD1100027916	11	96058994	0.2867
934	BovineHD1100028070	11	96654864	0.4946
935	BovineHD1100029093	11	100150378	0.3189
936	Hapmap23812-BTA-127135	11	100607031	0.3940
937	BovineHD1100029565	11	101648927	0.3822
938	BovineHD1100029628	11	101858112	0.4025
939	BovineHD1100029646	11	101918593	0.4234
940	BovineHD1100030721	11	106272264	0.4283
941	BovineHD1100030629	11	106695153	0.2342
942	BovineHD1200000065	12	525181	0.3251
943	BovineHD1200000277	12	1272214	0.4393
944	BovineHD1200001241	12	4091486	0.4504
945	BovineHD1200001864	12	6613549	0.4649
946	BovineHD1200002328	12	8365004	0.4372
947	BovineHD1200002530	12	9178316	0.4235
948	BovineHD1200003031	12	10691956	0.4347
949	BovineHD1200003675	12	12712812	0.3562

950	BovineHD1200004283	12	14368414	0.2704
951	ARS-BFGL-NGS-18629	12	15004714	0.4844
952	BovineHD1200004629	12	15462742	0.4040
953	BovineHD1200004883	12	16174096	0.4375
954	BovineHD1200005155	12	16910729	0.3508
955	BovineHD1200005475	12	18020287	0.4705
956	Hapmap49780-BTA-86092	12	19311292	0.3113
957	BovineHD1200005929	12	19561951	0.4075
958	BovineHD1200007083	12	23491088	0.4940
959	BovineHD1200007706	12	25481996	0.3707
960	BovineHD1200008209	12	27517798	0.4127
961	BovineHD1200008921	12	30277788	0.2803
962	BovineHD1200009368	12	31814723	0.4976
963	BovineHD1200010352	12	35054878	0.1582
964	BovineHD1200026907	12	38096750	0.2807
965	BovineHD1200011320	12	39527109	0.2340
966	BovineHD1200011874	12	41798834	0.4709
967	BovineHD1200012076	12	42895170	0.3443
968	BovineHD1200012407	12	44398050	0.4425
969	Hapmap43513-BTA-21673	12	44835259	0.1793
970	BovineHD1200027933	12	46232946	0.3036
971	BovineHD1200013052	12	47120355	0.4490
972	BovineHD1200013177	12	47664103	0.3898
973	BovineHD1200013244	12	47950625	0.2410
974	BovineHD1200013387	12	48481227	0.3589
975	BovineHD1200013489	12	48802717	0.2707
976	BovineHD1200013639	12	49395285	0.2543
977	BovineHD1200013883	12	50252849	0.4804
978	BovineHD1200014081	12	50839569	0.4739
979	ARS-BFGL-NGS-90133	12	50942013	0.4897
980	BovineHD1200015377	12	55183878	0.4831
981	BovineHD1200016149	12	58436461	0.3582
982	BovineHD1200016321	12	59149429	0.4648
983	BTB-01012231	12	60947813	0.2090
984	BTB-01980499	12	60997744	0.1112
985	BovineHD1200017812	12	64664137	0.3968
986	BovineHD1200018115	12	65855722	0.4017
987	Hapmap38187-BTA-105082	12	68681354	0.1247
988	BTB-00500725	12	69110181	0.1945
989	BovineHD1200022080	12	73875542	0.4852
990	BovineHD1200022230	12	74215241	0.4401
991	BovineHD1200022713	12	75608674	0.2855
992	ARS-BFGL-NGS-115889	12	75844220	0.4671
993	ARS-BFGL-NGS-38451	12	78674429	0.2713
994	BovineHD1200023946	12	79003516	0.4427
995	BovineHD1200024120	12	79462412	0.3892

996	BovineHD1200024183	12	79637659	0.4414
997	BovineHD1200025224	12	82743078	0.4624
998	BovineHD1200025898	12	84570029	0.4910
999	BovineHD1200026185	12	85408245	0.2026
1000	BTA-92225-no-rs	13	460596	0.1386
1001	BovineHD1300000237	13	1258355	0.4436
1002	BovineHD1300000297	13	1451862	0.3157
1003	BovineHD1300000349	13	1605357	0.4839
1004	ARS-BFGL-NGS-102805	13	2775064	0.1558
1005	BovineHD1300000829	13	3142144	0.4123
1006	BovineHD1300000840	13	3230206	0.1493
1007	BovineHD1300001206	13	4755506	0.4473
1008	BovineHD1300001242	13	4908300	0.4642
1009	BovineHD1300001329	13	5186215	0.4080
1010	BovineHD1300001736	13	6614129	0.4333
1011	BovineHD1300001859	13	6964592	0.4945
1012	BovineHD1300002309	13	8596813	0.4273
1013	BovineHD1300003269	13	11783602	0.3450
1014	BovineHD1300003628	13	12791107	0.2858
1015	BovineHD1300003826	13	13282070	0.4390
1016	BovineHD1300004774	13	16620520	0.3908
1017	ARS-BFGL-NGS-112250	13	17532109	0.2472
1018	BovineHD1300005203	13	17978743	0.4337
1019	BovineHD1300005422	13	18654751	0.4609
1020	Hapmap35524-SCAFFOLD5373_40104	13	19710676	0.3892
1021	BovineHD1300006424	13	21683101	0.5000
1022	BovineHD1300007727	13	26401422	0.4146
1023	BovineHD1300008078	13	27536881	0.4952
1024	BovineHD1300008095	13	27576529	0.2966
1025	ARS-BFGL-NGS-108308	13	28717073	0.2440
1026	BovineHD1300009335	13	31584735	0.3365
1027	BovineHD1300009489	13	32258616	0.4647
1028	BovineHD1300009979	13	34017629	0.4285
1029	BovineHD1300010023	13	34237440	0.4431
1030	BovineHD1300010098	13	34457955	0.3306
1031	BovineHD1300010296	13	35149728	0.3360
1032	BovineHD1300010683	13	36606673	0.4695
1033	ARS-BFGL-NGS-28202	13	38053920	0.2693
1034	BovineHD1300011123	13	38285812	0.4210
1035	BovineHD1300011836	13	40570310	0.3802
1036	BovineHD1300013919	13	47284141	0.4876
1037	ARS-BFGL-NGS-38562	13	47813637	0.2213
1038	BovineHD1300014298	13	48799842	0.2026
1039	BovineHD1300024593	13	49790015	0.1214
1040	BovineHD1300014989	13	52463462	0.3299
1041	BovineHD1300015452	13	53981453	0.2800

1042	BovineHD1300015883	13	55354431	0.4927
1043	BovineHD1300016482	13	56900273	0.1183
1044	ARS-BFGL-NGS-115363	13	57744564	0.3412
1045	BovineHD1300017072	13	59045634	0.2809
1046	BovineHD1300017168	13	59384114	0.4166
1047	BovineHD1300017737	13	61449287	0.4802
1048	ARS-BFGL-NGS-107931	13	62925626	0.3435
1049	BovineHD1300018513	13	64502108	0.2351
1050	BovineHD1300018636	13	64980164	0.4455
1051	Hapmap38307-BTA-33222	13	65421742	0.1537
1052	BovineHD1300018874	13	65977310	0.2905
1053	BovineHD1300019101	13	66715848	0.3593
1054	BTA-33341-no-rs	13	66918432	0.4674
1055	ARS-BFGL-NGS-112094	13	68527135	0.1489
1056	ARS-BFGL-NGS-44322	13	71067907	0.2742
1057	ARS-BFGL-NGS-38677	13	71369326	0.3146
1058	BovineHD1300020829	13	71732892	0.4375
1059	BovineHD1300020961	13	72141821	0.4349
1060	BovineHD1300022125	13	75706396	0.2823
1061	BovineHD1300023394	13	80040151	0.4130
1062	ARS-BFGL-NGS-29997	13	81659411	0.4057
1063	BovineHD1300024262	13	82616878	0.4505
1064	BovineHD1300024499	13	83427114	0.3297
1065	Hapmap26072-BTC-065132	14	3120000	0.2676
1066	BovineHD1400001026	14	3488760	0.4238
1067	BovineHD1400002022	14	6425334	0.3698
1068	BovineHD1400002051	14	6539869	0.4388
1069	BovineHD1400002432	14	7795359	0.3052
1070	BovineHD4100011043	14	9382539	0.3628
1071	BovineHD1400003686	14	11626687	0.3906
1072	BovineHD1400003898	14	12406301	0.4798
1073	BovineHD1400004133	14	13200984	0.3299
1074	BovineHD1400005047	14	16114657	0.3190
1075	BovineHD1400005773	14	18613954	0.4812
1076	BovineHD1400005882	14	18948852	0.4312
1077	ARS-BFGL-NGS-28234	14	19710934	0.1281
1078	ARS-BFGL-NGS-106196	14	21764414	0.4127
1079	BovineHD1400007058	14	22678296	0.4459
1080	ARS-USMARC-Parent-DQ984826-rs29027559	14	26068764	0.2141
1081	BovineHD1400007949	14	26091788	0.3349
1082	BovineHD1400008667	14	28236317	0.3386
1083	BovineHD1400008982	14	29418107	0.3214
1084	ARS-BFGL-NGS-13139	14	30581043	0.2165
1085	BovineHD1400009525	14	31218288	0.2795
1086	ARS-BFGL-NGS-112227	14	32712341	0.3587

1087	BovineHD1400010348	14	33830236	0.4830
1088	Hapmap41742-BTA-107716	14	34727883	0.2617
1089	BovineHD1400024470	14	35237657	0.4439
1090	BovineHD1400010943	14	36100767	0.4427
1091	BovineHD1400011220	14	36983561	0.2779
1092	BovineHD1400012656	14	42628595	0.4923
1093	BovineHD1400012990	14	43839829	0.4541
1094	BovineHD1400013265	14	44717102	0.4614
1095	BovineHD1400013724	14	46183201	0.4674
1096	BovineHD1400014729	14	49722735	0.3305
1097	BovineHD1400014765	14	49880501	0.4037
1098	BovineHD1400015135	14	51660213	0.3342
1099	Hapmap27948-BTA-129299	14	55072975	0.1339
1100	BovineHD1400016324	14	56615955	0.3943
1101	ARS-BFGL-NGS-118834	14	57489048	0.2635
1102	BovineHD1400017383	14	60326250	0.3654
1103	BovineHD1400017600	14	61303451	0.2395
1104	BovineHD1400018612	14	64347051	0.4802
1105	BovineHD1400019151	14	66250730	0.3453
1106	ARS-BFGL-NGS-57879	14	66390973	0.2924
1107	BovineHD1400019294	14	66615463	0.2223
1108	BovineHD1400019729	14	67951815	0.4420
1109	BovineHD1400020408	14	70455780	0.3744
1110	ARS-BFGL-NGS-68191	14	70899906	0.3882
1111	BovineHD1400021458	14	74170064	0.3715
1112	BTB-00853404	14	75782925	0.0999
1113	BovineHD1400022257	14	76919153	0.2070
1114	BovineHD1400022640	14	78045230	0.4079
1115	BovineHD1400022834	14	78445219	0.2820
1116	BovineHD1400023102	14	79451202	0.1362
1117	BovineHD1400023340	14	80315157	0.2305
1118	BovineHD1500000073	15	438420	0.4321
1119	UA-IFASA-7453	15	2091409	0.1480
1120	BovineHD1500000729	15	3035577	0.2636
1121	BovineHD1500000798	15	3311669	0.3129
1122	BTB-01576968	15	4776153	0.4897
1123	BTB-01802172	15	4894029	0.4306
1124	Hapmap57381-rs29020495	15	5758635	0.2503
1125	BovineHD1500001598	15	6104661	0.3002
1126	BovineHD1500001641	15	6281169	0.4758
1127	BovineHD1500001856	15	7187724	0.3218
1128	BovineHD1500002290	15	8709242	0.4468
1129	BovineHD1000027411	15	11006803	0.3546
1130	BovineHD1500003025	15	11600771	0.4085
1131	BovineHD1500003237	15	12525222	0.4254
1132	BovineHD1500003951	15	15498665	0.4686

1133	BTB-01822806	15	16678746	0.2923
1134	BovineHD1500004546	15	17560627	0.1705
1135	ARS-BFGL-NGS-35704	15	18066576	0.1164
1136	BovineHD1500004712	15	18353428	0.2153
1137	BovineHD1500005275	15	20432391	0.4609
1138	ARS-BFGL-NGS-35405	15	21174615	0.2884
1139	BovineHD1500005526	15	21477568	0.3468
1140	BovineHD1500005845	15	22576694	0.4468
1141	BovineHD1500006311	15	23868659	0.1619
1142	BovineHD1500006312	15	23871012	0.0886
1143	ARS-BFGL-NGS-116737	15	26151739	0.0668
1144	BovineHD1500007174	15	26507947	0.3743
1145	BovineHD1500007537	15	27651159	0.4249
1146	BovineHD1500008619	15	31351181	0.4207
1147	BovineHD1500008735	15	31491359	0.3507
1148	BovineHD1500008850	15	32071580	0.2128
1149	BovineHD1500009559	15	34497030	0.2945
1150	ARS-BFGL-NGS-96067	15	35632101	0.2428
1151	BovineHD1500010047	15	36110147	0.4029
1152	ARS-BFGL-NGS-34422	15	40279882	0.2647
1153	BovineHD1500011393	15	41011469	0.4741
1154	BTB-01371672	15	49801112	0.2288
1155	ARS-BFGL-NGS-11811	15	50508537	0.3852
1156	ARS-BFGL-NGS-40224	15	53839563	0.4915
1157	ARS-BFGL-NGS-41239	15	55701529	0.2387
1158	BovineHD1500016840	15	57545603	0.3289
1159	ARS-BFGL-BAC-6489	15	57629970	0.1175
1160	BTB-00607669	15	58241375	0.1200
1161	BovineHD1500017391	15	59809005	0.4310
1162	BovineHD1500018362	15	63285089	0.4707
1163	BovineHD4100012144	15	64813742	0.2885
1164	BovineHD1500018934	15	65232445	0.1889
1165	BovineHD1500018950	15	65296117	0.2224
1166	BovineHD1500018983	15	65393674	0.1600
1167	BTB-01649233	15	66324113	0.3321
1168	BovineHD1500019469	15	66750160	0.4257
1169	BovineHD1500020670	15	70793346	0.4822
1170	BovineHD1500020981	15	71910961	0.4298
1171	BovineHD1500021178	15	72518013	0.4765
1172	ARS-BFGL-NGS-4967	15	73018714	0.4268
1173	BovineHD1500022411	15	75973861	0.4572
1174	BovineHD1500022891	15	77532464	0.4602
1175	BovineHD1500024279	15	81596439	0.4369
1176	BovineHD1500024471	15	82033564	0.1243
1177	ARS-BFGL-NGS-17641	16	756359	0.2806
1178	BovineHD1600024254	16	3996364	0.2220

1179	BovineHD1600024258	16	5223346	0.2427
1180	ARS-BFGL-NGS-89740	16	6329170	0.4012
1181	BovineHD1600002369	16	7767384	0.4146
1182	BovineHD1600002923	16	10296418	0.4981
1183	ARS-BFGL-NGS-114959	16	11003793	0.3795
1184	BovineHD1600003404	16	12170676	0.4899
1185	BovineHD1600003539	16	12703013	0.4209
1186	BovineHD1600003659	16	13051472	0.2937
1187	ARS-BFGL-BAC-19454	16	13072437	0.1535
1188	BovineHD1600004079	16	14412122	0.3770
1189	Hapmap42539-BTA-40292	16	16433701	0.1875
1190	Hapmap50501-BTA-91866	16	17824360	0.1841
1191	BovineHD1600005161	16	18361905	0.3293
1192	Hapmap50635-BTA-26852	16	18835344	0.3879
1193	Hapmap42200-BTA-40314	16	19031403	0.3079
1194	Hapmap52022-BTA-40336	16	20087389	0.0997
1195	Hapmap60688-rs29014090	16	20467205	0.3500
1196	ARS-BFGL-NGS-111618	16	21463795	0.1704
1197	BovineHD1600006334	16	21832324	0.3441
1198	BovineHD1600006629	16	23238995	0.4591
1199	BovineHD1600006657	16	23372137	0.4984
1200	BovineHD1600007032	16	24458700	0.2297
1201	BovineHD1600007140	16	24944651	0.4159
1202	BovineHD1600007196	16	25180364	0.4852
1203	BovineHD1600024421	16	25194244	0.4891
1204	ARS-BFGL-NGS-54489	16	27938714	0.1657
1205	BovineHD1600008650	16	29745892	0.4731
1206	BovineHD4100012453	16	32456121	0.3676
1207	BovineHD1600009851	16	33655657	0.2513
1208	BovineHD1600009861	16	33669683	0.4017
1209	BovineHD1600010027	16	34216764	0.4068
1210	BovineHD1600010167	16	34614526	0.4917
1211	BovineHD1600010360	16	35269218	0.3976
1212	ARS-BFGL-NGS-41002	16	35613606	0.1772
1213	ARS-BFGL-NGS-39390	16	37251454	0.2322
1214	BovineHD1600011883	16	41716228	0.4716
1215	BovineHD1600024521	16	42405866	0.4691
1216	BovineHD1600012340	16	43403106	0.4220
1217	BovineHD1600012680	16	45103219	0.4029
1218	BovineHD1600012742	16	45430353	0.2791
1219	BovineHD1600013019	16	46279513	0.1180
1220	BovineHD1600013526	16	47952985	0.3130
1221	ARS-BFGL-NGS-32604	16	50274869	0.2568
1222	BovineHD1600014631	16	51721240	0.4345
1223	BovineHD1600014789	16	52140360	0.2049
1224	BovineHD1600015464	16	54084030	0.2694

1225	BovineHD1600015599	16	54614265	0.0796
1226	BovineHD1600016046	16	56164143	0.2760
1227	BovineHD1600016738	16	58412746	0.1648
1228	BovineHD1600017354	16	60352621	0.3958
1229	Hapmap25615-BTA-160036	16	62068790	0.3538
1230	BovineHD1600017962	16	62455462	0.4809
1231	BovineHD1600018636	16	64354689	0.4936
1232	BovineHD1600019043	16	65631580	0.3767
1233	Hapmap59847-rs29025534	16	66859242	0.1057
1234	Hapmap50200-BTA-106536	16	69857792	0.4771
1235	Hapmap47281-BTA-40051	16	71035799	0.2908
1236	BovineHD1600020758	16	71151975	0.4957
1237	BovineHD1600021436	16	73068141	0.3570
1238	BovineHD1600021613	16	73529151	0.4953
1239	BovineHD1600021789	16	73889247	0.4809
1240	BovineHD4100012745	16	74377168	0.3108
1241	BovineHD1600022732	16	76312874	0.1892
1242	BovineHD1600022829	16	76545440	0.3402
1243	BovineHD1600023350	16	78142248	0.3132
1244	BTA-95562-no-rs	16	78483462	0.3458
1245	ARS-BFGL-NGS-52551	16	78936848	0.4941
1246	BovineHD1700000359	17	1638365	0.4537
1247	BovineHD1700021903	17	3148158	0.4748
1248	BovineHD1700001442	17	5411321	0.4140
1249	ARS-BFGL-NGS-38059	17	5843870	0.1395
1250	BovineHD1700001996	17	6923401	0.4248
1251	BovineHD1700002126	17	7340739	0.3783
1252	BovineHD1700002743	17	9641431	0.4149
1253	BovineHD1700003297	17	11342602	0.1429
1254	BovineHD1700003889	17	13478984	0.4814
1255	BovineHD1700004866	17	16679812	0.2693
1256	BovineHD1700005020	17	17078526	0.4672
1257	ARS-BFGL-NGS-28104	17	17128570	0.3079
1258	BovineHD1700005215	17	17707083	0.4317
1259	BovineHD1700006138	17	21003064	0.3825
1260	BovineHD1700006210	17	21276795	0.1709
1261	BovineHD1700006743	17	23279150	0.4339
1262	BovineHD1700007286	17	25426680	0.4683
1263	BovineHD1700007768	17	27047686	0.4165
1264	BovineHD1700009716	17	34705389	0.4299
1265	BovineHD1700009832	17	35180086	0.4628
1266	BovineHD1700010002	17	35935403	0.4878
1267	BovineHD1700010149	17	36348625	0.4619
1268	BovineHD1700011078	17	39551347	0.4569
1269	ARS-BFGL-NGS-1481	17	43283637	0.4821
1270	BovineHD1700013420	17	46792851	0.3871

1271	BovineHD1700014135	17	48722377	0.3730
1272	ARS-BFGL-NGS-99210	17	50485602	0.3525
1273	BovineHD1700015340	17	51942830	0.4868
1274	ARS-BFGL-NGS-21400	17	54862080	0.2712
1275	BovineHD1700016337	17	55368214	0.3666
1276	BovineHD1700016875	17	57300083	0.4352
1277	BovineHD1700017421	17	58902969	0.2690
1278	BovineHD1700017556	17	59332427	0.4895
1279	BovineHD1700017770	17	59987188	0.3376
1280	ARS-BFGL-NGS-39993	17	60781502	0.2931
1281	BovineHD1700018450	17	61645564	0.4706
1282	BovineHD1700019552	17	65243374	0.4871
1283	BovineHD1700019881	17	66080615	0.3002
1284	BovineHD1700020443	17	68002669	0.3642
1285	BovineHD1700022145	17	68962935	0.3449
1286	BovineHD1700021195	17	70515609	0.4509
1287	BovineHD1700021263	17	70644195	0.2556
1288	BovineHD1700021431	17	71355374	0.2969
1289	BovineHD1700021714	17	72272981	0.4450
1290	BovineHD1800000105	18	840728	0.2929
1291	ARS-BFGL-NGS-21711	18	1113034	0.3942
1292	ARS-USMARC-Parent-EF028073-rs29014953	18	1801945	0.2028
1293	BovineHD1800000625	18	2259433	0.3381
1294	BovineHD1800000707	18	2642750	0.3199
1295	BTB-01040984	18	3039492	0.3946
1296	Hapmap54020-rs29023153	18	6747828	0.2366
1297	ARS-BFGL-NGS-111796	18	9607032	0.4559
1298	BovineHD1800004333	18	12959558	0.4438
1299	BovineHD1800019530	18	14504736	0.3828
1300	ARS-BFGL-NGS-112592	18	18185722	0.3066
1301	BovineHD1800006001	18	19199482	0.3907
1302	ARS-BFGL-NGS-98964	18	19318313	0.3415
1303	BovineHD1800006239	18	20072380	0.3809
1304	BovineHD1800006823	18	22102391	0.2682
1305	BovineHD1800007167	18	23344694	0.4605
1306	BovineHD1800007513	18	24230415	0.4689
1307	BovineHD1800007648	18	24729491	0.4501
1308	ARS-BFGL-NGS-19431	18	28396317	0.3372
1309	BovineHD1800009036	18	29036484	0.3849
1310	BovineHD1800009658	18	31810388	0.4828
1311	BovineHD1800009971	18	32828541	0.4593
1312	ARS-BFGL-NGS-112014	18	34674118	0.3772
1313	BovineHD1800011162	18	37395898	0.4650
1314	BTB-00711650	18	37743151	0.3733
1315	BovineHD1800011891	18	40236316	0.3870
1316	BovineHD1800019683	18	40961440	0.4851

1317	BovineHD1800012555	18	42339745	0.2225
1318	BovineHD1800012855	18	43310489	0.4851
1319	BovineHD1800013673	18	46307516	0.3843
1320	BovineHD1800013905	18	46947848	0.4157
1321	ARS-USMARC-Parent-AY914316-rs17871403	18	48545897	0.4520
1322	BTA-43602-no-rs	18	48840710	0.4955
1323	BovineHD1800014626	18	49388468	0.4620
1324	BovineHD1800015260	18	51704629	0.3722
1325	BovineHD1800015530	18	52364861	0.2450
1326	BovineHD1800019789	18	53584535	0.4547
1327	BovineHD1800016248	18	55052084	0.2855
1328	BovineHD1800016360	18	55454395	0.3812
1329	ARS-BFGL-NGS-71741	18	56679475	0.1565
1330	BovineHD1800018741	18	64475993	0.4350
1331	BovineHD1800018922	18	64937253	0.4545
1332	BovineHD1800018967	18	65001458	0.3595
1333	BovineHD1800019037	18	65179636	0.4917
1334	BovineHD1800019059	18	65261044	0.1466
1335	ARS-BFGL-NGS-92544	19	1124406	0.3078
1336	BovineHD4100013836	19	1695784	0.2762
1337	BovineHD1900000465	19	1978988	0.4451
1338	Hapmap50598-BTA-122724	19	4478391	0.3797
1339	BovineHD1900001402	19	5431826	0.4344
1340	BovineHD1900001572	19	5970293	0.4152
1341	Hapmap51233-BTA-44716	19	7065374	0.3318
1342	BovineHD1900001941	19	7133600	0.3859
1343	BovineHD1900018767	19	9610225	0.4836
1344	BovineHD1900003109	19	11335930	0.1106
1345	BovineHD1900003121	19	11430581	0.4791
1346	BovineHD1900003381	19	12408624	0.3357
1347	BovineHD1900003500	19	12981194	0.4473
1348	ARS-USMARC-Parent-EF026084-rs29025380	19	15017885	0.1605
1349	ARS-BFGL-NGS-112774	19	15653968	0.1468
1350	BovineHD1900004513	19	16064959	0.4548
1351	BovineHD1900004781	19	16787943	0.4250
1352	UA-IFASA-7987	19	18449930	0.2206
1353	BovineHD1900005466	19	18814313	0.3572
1354	BovineHD1900006263	19	21301396	0.4913
1355	BovineHD1900006310	19	21531680	0.3928
1356	BovineHD1900006346	19	21714004	0.4279
1357	BovineHD1900007618	19	25293021	0.2821
1358	BovineHD1900007686	19	25439551	0.3402
1359	BovineHD1900008733	19	29102715	0.4656
1360	BovineHD1900008854	19	29463611	0.3824
1361	Hapmap59534-rs29022537	19	31547330	0.3551

1362	BovineHD1900009485	19	31591985	0.4854
1363	BovineHD1900018886	19	32797444	0.3763
1364	BovineHD1900010366	19	35014810	0.3652
1365	BovineHD1900018924	19	38040560	0.2933
1366	Hapmap28098-BTA-150233	19	38279949	0.4573
1367	BovineHD1900011616	19	40162636	0.4973
1368	BovineHD1900011727	19	40543571	0.0902
1369	BovineHD1900011787	19	40744460	0.3628
1370	BovineHD1900011825	19	40857061	0.4044
1371	BovineHD1900012142	19	41925621	0.1462
1372	ARS-BFGL-NGS-22409	19	42655200	0.3229
1373	BovineHD1900012665	19	44451884	0.4636
1374	BovineHD1900012957	19	45676351	0.3324
1375	BovineHD1900013578	19	48167608	0.2819
1376	BovineHD1900014341	19	50675382	0.3713
1377	ARS-BFGL-NGS-114192	19	53386469	0.0701
1378	BovineHD1900015542	19	54597794	0.2777
1379	BovineHD1900015968	19	55864707	0.3883
1380	BovineHD1900015989	19	55917441	0.2217
1381	BovineHD1900016011	19	55988158	0.4610
1382	Hapmap48769-BTA-46039	19	56224509	0.3001
1383	BovineHD1900016521	19	57523884	0.3434
1384	BovineHD1900017636	19	60784762	0.4976
1385	BovineHD1900017846	19	61423018	0.2770
1386	ARS-USMARC-Parent-DQ984828-rs29010004	20	788373	0.1113
1387	BovineHD2000000740	20	2171874	0.2526
1388	Hapmap36588-SCAFFOLD90561_9460	20	2386360	0.3583
1389	BovineHD2000000997	20	3048441	0.4035
1390	BovineHD2000001051	20	3238459	0.2022
1391	Hapmap53870-rs29020081	20	6076915	0.4518
1392	BovineHD2000002168	20	6987225	0.3965
1393	BovineHD2000002650	20	8354113	0.4991
1394	BovineHD2000002946	20	9398783	0.4270
1395	BovineHD2000003554	20	11212305	0.4734
1396	ARS-BFGL-NGS-84611	20	11948110	0.3380
1397	BovineHD2000004513	20	14613502	0.1571
1398	BovineHD2000004643	20	15422484	0.4258
1399	BovineHD2000004729	20	15681811	0.3621
1400	BovineHD2000004845	20	16130995	0.4873
1401	BovineHD2000006226	20	20798417	0.4028
1402	BovineHD2000006493	20	21676211	0.1242
1403	BovineHD2000006624	20	22122113	0.2950
1404	BovineHD2000006683	20	22228621	0.4424
1405	BovineHD2000006703	20	22300653	0.2011
1406	BovineHD2000006742	20	22418921	0.2353

1407	BovineHD2000007029	20	23332211	0.4647
1408	BovineHD2000021292	20	26774554	0.3299
1409	BovineHD2000008464	20	28607857	0.4481
1410	BovineHD2000008935	20	30671790	0.4641
1411	BovineHD2000009358	20	32680643	0.1667
1412	BovineHD2000009613	20	33451856	0.4714
1413	BovineHD2000010409	20	36290830	0.4622
1414	BovineHD2000010665	20	37225471	0.4977
1415	BovineHD2000011217	20	39364262	0.3845
1416	ARS-BFGL-NGS-114097	20	41684531	0.1247
1417	Hapmap35044-BES8_Contig309_741	20	42113227	0.2968
1418	BovineHD2000012405	20	43519291	0.2535
1419	Hapmap50991-BTA-50645	20	44280089	0.1373
1420	BTA-50676-no-rs	20	44673055	0.2998
1421	Hapmap35695-SCAFFOLD130061_3178	20	45459536	0.1612
1422	ARS-USMARC-Parent-AY919868-rs29002211	20	46043225	0.2387
1423	BovineHD2000014019	20	50582788	0.3013
1424	BovineHD2000014287	20	51688767	0.3738
1425	BTB-02040655	20	52212800	0.2487
1426	BovineHD2000015047	20	54827248	0.4911
1427	BovineHD2000015167	20	55315101	0.3949
1428	BovineHD2000015301	20	55806033	0.3273
1429	ARS-BFGL-NGS-111474	20	57804168	0.3399
1430	BovineHD2000016774	20	59605319	0.4858
1431	BovineHD2000016934	20	60269475	0.4025
1432	BovineHD2000017791	20	62686569	0.3700
1433	ARS-BFGL-NGS-110376	20	63620380	0.4757
1434	ARS-BFGL-NGS-41833	20	66360651	0.3555
1435	BovineHD2000019550	20	67264220	0.4243
1436	BovineHD2000020795	20	70971513	0.4891
1437	BovineHD2100000046	21	1512376	0.4426
1438	BovineHD2100000008	21	1920814	0.4605
1439	BovineHD2100000478	21	3107259	0.4905
1440	BovineHD2100000930	21	5108764	0.2233
1441	BovineHD2100001401	21	6629678	0.4331
1442	BovineHD2100001494	21	6883116	0.4914
1443	ARS-BFGL-BAC-34544	21	7872837	0.3030
1444	BovineHD2100002236	21	9260887	0.4218
1445	BovineHD2100002411	21	9792846	0.4947
1446	BovineHD2100002908	21	11464874	0.3456
1447	BovineHD2100002947	21	11623713	0.3820
1448	ARS-BFGL-NGS-109907	21	12296006	0.1249
1449	BovineHD2100003994	21	14687345	0.4906
1450	BovineHD2100004170	21	15044963	0.4147
1451	BovineHD2100005029	21	17474484	0.3816

1452	Hapmap57137-rs29011525	21	18086696	0.1776
1453	ARS-BFGL-NGS-111076	21	20809473	0.2106
1454	BovineHD2100006394	21	21279682	0.4535
1455	BovineHD2100006874	21	22953407	0.4994
1456	BovineHD2100007092	21	23828481	0.4455
1457	Hapmap39461-BTA-109898	21	27418010	0.0966
1458	BovineHD2100008824	21	30141433	0.4378
1459	BovineHD2100008867	21	30280966	0.2349
1460	ARS-BFGL-NGS-117619	21	30510352	0.2888
1461	BovineHD2100021269	21	31899339	0.4043
1462	Hapmap25541-BTA-135872	21	32199683	0.2335
1463	BovineHD2100009676	21	33439800	0.4916
1464	BovineHD2100010257	21	35056505	0.4783
1465	BovineHD2100010360	21	35232116	0.3146
1466	BTB-02072074	21	37826412	0.4701
1467	BTB-00818821	21	39994586	0.1521
1468	BovineHD2100011764	21	40797255	0.3651
1469	BovineHD2100012139	21	42118649	0.4620
1470	ARS-BFGL-NGS-23813	21	42120151	0.2141
1471	BovineHD2100012222	21	42459276	0.4844
1472	Hapmap47652-BTA-52437	21	46779702	0.3006
1473	BTB-01686668	21	48237392	0.0976
1474	BovineHD2100014333	21	49381260	0.2546
1475	BovineHD2100015051	21	51888446	0.2738
1476	BovineHD2100015178	21	52331051	0.3866
1477	BovineHD2100016012	21	55365727	0.1883
1478	BovineHD2100016560	21	57032781	0.4566
1479	BovineHD2100016571	21	57047736	0.0632
1480	BovineHD2100016591	21	57116691	0.3414
1481	BovineHD2100016603	21	57144290	0.4552
1482	ARS-BFGL-NGS-69497	21	59541766	0.3433
1483	Hapmap31098-BTA-136127	21	60444453	0.1742
1484	ARS-BFGL-BAC-34585	21	60659885	0.1089
1485	ARS-USMARC-Parent-EF026085-rs29021607	21	63550319	0.4091
1486	BTA-52957-no-rs	21	64323158	0.1929
1487	Hapmap35472-SCAFFOLD35342_21388	21	64458034	0.2878
1488	BovineHD2100019895	21	66258428	0.4655
1489	BovineHD2100021436	21	67379846	0.3177
1490	BovineHD2100020508	21	68533964	0.4287
1491	ARS-BFGL-NGS-14557	22	1654768	0.2054
1492	BovineHD2200000658	22	2550086	0.3595
1493	BovineHD2200001013	22	3742292	0.4403
1494	BovineHD2200001079	22	3955125	0.4971
1495	BovineHD2200001206	22	4455421	0.2913
1496	ARS-BFGL-NGS-100995	22	5061438	0.4707
1497	BovineHD2200001710	22	5784520	0.4651

1498	BovineHD2200001775	22	5981628	0.3859
1499	BovineHD2200002123	22	6919584	0.3853
1500	Hapmap44095-BTA-55125	22	8558220	0.2129
1501	ARS-USMARC-Parent-DQ990832-rs29015065	22	11000418	0.0801
1502	BovineHD2200003310	22	11133772	0.4568
1503	BovineHD2200003402	22	11398960	0.0912
1504	Hapmap44181-BTA-104938	22	13473162	0.3291
1505	BovineHD2200004144	22	14178689	0.3790
1506	BovineHD2200004496	22	14683694	0.4889
1507	BovineHD2200004741	22	16361137	0.4664
1508	BovineHD2200004907	22	16893878	0.1115
1509	BovineHD2200004981	22	17274423	0.1464
1510	BovineHD2200005512	22	18898450	0.4466
1511	BovineHD2200005981	22	20649707	0.1913
1512	BovineHD2200006097	22	20998263	0.4366
1513	UA-IFASA-6532	22	21678074	0.2651
1514	ARS-USMARC-Parent-EF093509-rs29015170	22	22468414	0.1025
1515	BTA-16900-no-rs	22	23400462	0.4158
1516	BTA-53914-no-rs	22	24946405	0.3855
1517	BovineHD2200007537	22	25739885	0.3041
1518	BovineHD2200007798	22	27065087	0.4958
1519	BovineHD2200008116	22	27953380	0.4293
1520	BovineHD2200008662	22	29802566	0.3900
1521	BovineHD2200008803	22	30452058	0.3742
1522	BovineHD2200010327	22	35966718	0.4071
1523	BovineHD2200010828	22	37884377	0.2458
1524	BovineHD2200011085	22	38690293	0.3569
1525	BovineHD2200011320	22	39372716	0.2077
1526	BovineHD2200011380	22	39535415	0.4523
1527	BovineHD2200011510	22	39906000	0.4869
1528	BovineHD2200011740	22	40533527	0.4402
1529	BovineHD2200012366	22	42138172	0.4126
1530	BovineHD2200012572	22	42853528	0.4819
1531	BovineHD2200013566	22	46688927	0.3531
1532	ARS-BFGL-NGS-108083	22	46981263	0.1797
1533	BovineHD2200013690	22	47031388	0.4984
1534	BovineHD2200014033	22	48353575	0.4536
1535	Hapmap47041-BTA-54700	22	48924379	0.4299
1536	BovineHD2200015286	22	53330863	0.4880
1537	BovineHD2200016184	22	55741150	0.4219
1538	ARS-BFGL-NGS-102750	22	56963579	0.0900
1539	BovineHD2200017033	22	58134032	0.4122
1540	BovineHD2200017108	22	58355582	0.4843
1541	BovineHD2200017767	22	60333848	0.3434
1542	BovineHD2300000550	23	2945377	0.3825

1543	ARS-BFGL-NGS-115177	23	7930555	0.2412
1544	BovineHD2300015250	23	8047326	0.4092
1545	BovineHD2300015260	23	9932862	0.4240
1546	BovineHD2300003011	23	12129871	0.3413
1547	BovineHD2300003384	23	13562311	0.3772
1548	ARS-BFGL-NGS-34042	23	16043022	0.1315
1549	BovineHD2300015569	23	16318054	0.3233
1550	BovineHD2300004246	23	16521600	0.2139
1551	ARS-BFGL-NGS-10525	23	17944024	0.4691
1552	ARS-BFGL-NGS-117343	23	20274918	0.4192
1553	BovineHD2300006061	23	23299535	0.4207
1554	BovineHD2300006067	23	23309186	0.3893
1555	BovineHD2300006663	23	24825970	0.4581
1556	BovineHD2300008428	23	29757806	0.2822
1557	Hapmap36280-SCAFFOLD155216_10397	23	30176828	0.1042
1558	BovineHD2300008652	23	30404637	0.3331
1559	BovineHD2300008951	23	31293488	0.4357
1560	BovineHD2300009081	23	31683518	0.4999
1561	Hapmap59614-ss46526680	23	33181852	0.4877
1562	BovineHD2300009827	23	33718755	0.4544
1563	BovineHD2300009914	23	33980949	0.2366
1564	BovineHD2300015668	23	35602977	0.3733
1565	BovineHD2300010747	23	37359163	0.1486
1566	BovineHD2300010763	23	37424746	0.2777
1567	BovineHD2300011338	23	39459080	0.2086
1568	BovineHD2300011346	23	39492347	0.4333
1569	BovineHD2300011359	23	39542549	0.2185
1570	BovineHD2300015701	23	39975879	0.4071
1571	BovineHD2300012320	23	41335957	0.4429
1572	BovineHD2300012074	23	42267827	0.3784
1573	BovineHD2300012153	23	42506697	0.4193
1574	BovineHD4100016228	23	43199690	0.4961
1575	BovineHD2300012648	23	43706222	0.4020
1576	Hapmap38722-BTA-107506	23	45646476	0.1335
1577	BovineHD2300013381	23	46218062	0.3052
1578	BovineHD2300015736	23	46454950	0.4686
1579	BTB-00869068	23	47297677	0.1830
1580	BovineHD2300013831	23	47847042	0.2743
1581	BovineHD2300014965	23	51604405	0.4234
1582	BovineHD2300015082	23	52044193	0.3510
1583	BovineHD4100016347	24	1063713	0.3752
1584	BovineHD2400000656	24	2232038	0.4258
1585	BovineHD2400001367	24	4651334	0.1819
1586	Hapmap34424-BES10_Contig566_926	24	6125588	0.4343
1587	BTB-01910403	24	7918389	0.3000
1588	BovineHD2400002492	24	8488923	0.4929

1589	BovineHD2400003154	24	10854635	0.2827
1590	BovineHD2400003328	24	11541673	0.3507
1591	ARS-USMARC-Parent-DQ990833-rs29010147	24	15141986	0.3764
1592	ARS-BFGL-NGS-14815	24	16464296	0.4311
1593	Hapmap42592-BTA-57478	24	16874416	0.4809
1594	BovineHD2400004922	24	18370828	0.4453
1595	BovineHD2400005210	24	19398902	0.4936
1596	BovineHD2400005258	24	19559620	0.4226
1597	ARS-BFGL-NGS-73114	24	19851885	0.1620
1598	BovineHD2400005891	24	21478537	0.4655
1599	BTA-57636-no-rs	24	22087989	0.3665
1600	BovineHD2400006230	24	22497294	0.3683
1601	BovineHD2400006505	24	23533568	0.3225
1602	BovineHD2400007146	24	25959122	0.3738
1603	Hapmap34487-BES10_Contig784_2526	24	26680497	0.2755
1604	BovineHD2400007394	24	26717138	0.4564
1605	BTA-57747-no-rs	24	28244099	0.3068
1606	BovineHD2400008103	24	29491437	0.3660
1607	BovineHD2400008139	24	29694735	0.4223
1608	BovineHD2400008644	24	31302891	0.4861
1609	BovineHD2400008669	24	31512043	0.2841
1610	BovineHD2400009314	24	33287794	0.4286
1611	BovineHD2400009966	24	35948503	0.3394
1612	BovineHD2400010066	24	36444258	0.3663
1613	BovineHD2400010149	24	36736535	0.3197
1614	BovineHD2400010281	24	37200329	0.4422
1615	BovineHD2400010536	24	38037557	0.4331
1616	BovineHD2400011648	24	41532521	0.4803
1617	ARS-BFGL-NGS-32328	24	44375642	0.1547
1618	BovineHD2400012561	24	45282908	0.3902
1619	Hapmap38711-BTA-90355	24	45948420	0.2807
1620	BovineHD2400013036	24	46469125	0.3920
1621	Hapmap52259-rs29019756	24	47155098	0.2268
1622	BovineHD2400013839	24	49095913	0.3790
1623	BovineHD2400014032	24	49789624	0.3849
1624	ARS-BFGL-NGS-118573	24	51130282	0.1540
1625	BovineHD2400014662	24	51594960	0.4540
1626	BTA-117811-no-rs	24	53836923	0.1712
1627	BovineHD2400015590	24	54162021	0.4330
1628	ARS-BFGL-NGS-35634	24	55513422	0.1488
1629	BovineHD2400016579	24	57348868	0.4070
1630	BovineHD2400016949	24	58336565	0.4838
1631	ARS-BFGL-NGS-73260	24	58671488	0.1987
1632	BovineHD2400017123	24	58815843	0.4918
1633	BovineHD2400017286	24	59340107	0.4937

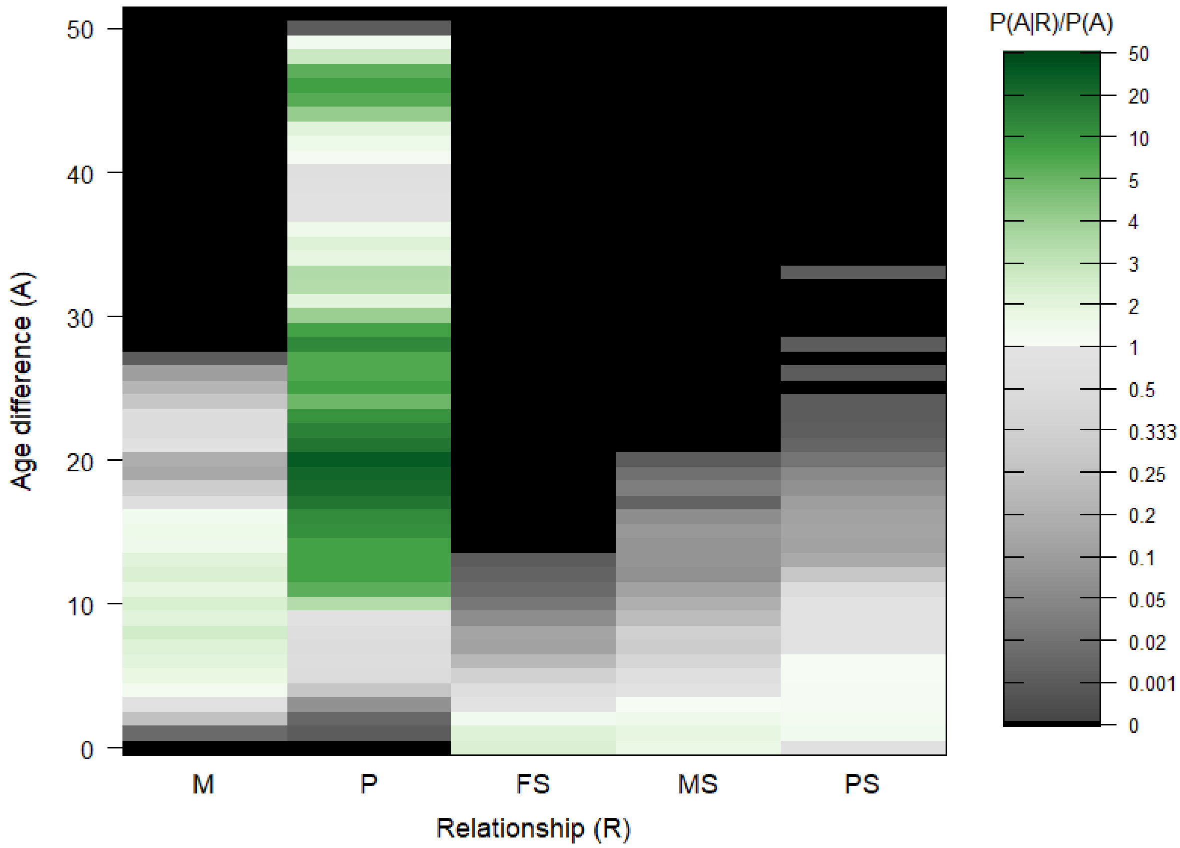
1634	BovineHD2400017913	24	61155084	0.4691
1635	BovineHD2500012112	25	191082	0.4554
1636	ARS-BFGL-NGS-19158	25	428640	0.2900
1637	BovineHD2500000518	25	2469673	0.4815
1638	Hapmap26816-BTC-015808	25	2777288	0.2074
1639	BovineHD2500000698	25	3223240	0.4275
1640	BovineHD2500000918	25	4228461	0.2568
1641	BovineHD2500001110	25	4778712	0.3622
1642	BovineHD2500012136	25	6165542	0.1504
1643	BovineHD2500002721	25	10156805	0.4859
1644	BovineHD2500003290	25	11748095	0.2354
1645	BovineHD2500003351	25	11906314	0.3541
1646	BovineHD2500003436	25	12169266	0.2720
1647	BovineHD2500003446	25	12212363	0.4200
1648	BovineHD2500004047	25	14305593	0.3666
1649	BovineHD2500004348	25	15318541	0.4139
1650	BovineHD2500004600	25	16170042	0.4572
1651	ARS-BFGL-NGS-115486	25	17057923	0.4837
1652	BovineHD2500005097	25	17988884	0.4848
1653	BovineHD2500005241	25	18516556	0.4957
1654	BovineHD2500005807	25	20791530	0.3458
1655	BovineHD2500005935	25	21174119	0.4875
1656	Hapmap50415-BTA-59713	25	23307044	0.4794
1657	BovineHD2500007904	25	28028129	0.1778
1658	BovineHD2500007949	25	28330891	0.4834
1659	BovineHD2500008087	25	28795501	0.0758
1660	Hapmap23570-BTC-029699	25	29964743	0.4650
1661	Hapmap25329-BTA-159835	25	31569710	0.4563
1662	BovineHD2500009460	25	33809534	0.4105
1663	BovineHD2500010041	25	35776120	0.2783
1664	BovineHD2500010153	25	36177124	0.1854
1665	BovineHD2500010238	25	36441453	0.0858
1666	BovineHD2500010554	25	37368155	0.4325
1667	BovineHD2500010731	25	37966187	0.0607
1668	ARS-BFGL-NGS-60739	25	38952307	0.4426
1669	BovineHD2500011964	25	42211471	0.4073
1670	BTB-01735218	26	1202126	0.4544
1671	BovineHD2600000913	26	4617733	0.4488
1672	BovineHD2600001328	26	5893198	0.4153
1673	ARS-BFGL-NGS-98526	26	6693744	0.0650
1674	BovineHD2600001673	26	7239097	0.3312
1675	BovineHD2600001731	26	7448946	0.4768
1676	BovineHD2600001997	26	8088819	0.3458
1677	BovineHD2600015008	26	10656674	0.4601
1678	BTA-27150-no-rs	26	12409311	0.1712
1679	BovineHD2600003427	26	13181206	0.2452

1680	BovineHD2600003455	26	13296326	0.4953
1681	BovineHD2600003750	26	14627405	0.3817
1682	BovineHD2600004906	26	19065632	0.3921
1683	BovineHD2600014958	26	20802996	0.4990
1684	BTB-00931586	26	21539987	0.0885
1685	BovineHD2600005581	26	21696804	0.2831
1686	BovineHD2600015168	26	23034520	0.4535
1687	BovineHD2600006348	26	24391655	0.4801
1688	BovineHD2600006753	26	25467889	0.2512
1689	BovineHD2600007904	26	29327785	0.4110
1690	BovineHD2600008325	26	30918659	0.2771
1691	BovineHD2600008468	26	31356117	0.4927
1692	Hapmap48021-BTA-61134	26	31647548	0.2681
1693	BovineHD2600009703	26	35225543	0.4957
1694	BovineHD2600009881	26	35844121	0.4987
1695	BovineHD2600010452	26	37670578	0.3288
1696	ARS-USMARC-Parent-EF034086-no-rs	26	37900334	0.4703
1697	BovineHD2600010638	26	38531298	0.1870
1698	Hapmap42736-BTA-93531	26	39030371	0.2871
1699	BovineHD2600011259	26	40460199	0.3780
1700	BTA-52296-no-rs	26	43197866	0.2572
1701	ARS-BFGL-NGS-30392	26	44205994	0.3182
1702	BTA-61758-no-rs	26	45008042	0.1505
1703	BovineHD2600013369	26	46546694	0.3885
1704	ARS-BFGL-NGS-74523	26	46602209	0.1821
1705	BovineHD2600013832	26	47691216	0.4017
1706	BovineHD2600014616	26	50103841	0.4523
1707	ARS-BFGL-NGS-117590	26	51824417	0.2130
1708	BovineHD2700000020	27	1223920	0.3004
1709	BovineHD2700000115	27	1451020	0.4655
1710	BovineHD2700000702	27	3193257	0.4126
1711	BovineHD2700001264	27	4838149	0.4623
1712	ARS-BFGL-NGS-116642	27	6073297	0.2351
1713	BovineHD2700002025	27	7539593	0.4565
1714	BovineHD2700002692	27	9669369	0.2787
1715	BovineHD2700002997	27	10990341	0.3928
1716	BovineHD2700003841	27	14109933	0.3311
1717	BTB-00623969	27	17989233	0.3715
1718	BTB-00000948	27	18223997	0.1083
1719	BovineHD2700005291	27	19462450	0.4684
1720	BTB-02036330	27	22507389	0.3818
1721	BTA-66900-no-rs	27	23304183	0.4959
1722	Hapmap54743-rs29013433	27	24812065	0.1161
1723	BovineHD2700007018	27	25830932	0.4312
1724	BovineHD2700007413	27	27306154	0.3085
1725	BovineHD2700013305	27	29788997	0.1981

1726	BovineHD2700008154	27	29931904	0.4916
1727	BovineHD2700008983	27	32271147	0.4702
1728	BovineHD2700009003	27	32313102	0.3888
1729	BovineHD2700009229	27	33000330	0.4268
1730	BovineHD2700013640	27	37233513	0.4924
1731	ARS-USMARC-Parent-EF141102-rs29015783	27	37824983	0.1628
1732	BovineHD2700010932	27	38273045	0.4030
1733	BovineHD2700011192	27	38860509	0.2377
1734	ARS-BFGL-NGS-76581	27	39928560	0.2910
1735	BovineHD2700011589	27	40117781	0.4009
1736	Hapmap51908-BTA-63031	27	41177956	0.2186
1737	BovineHD2700012739	27	44069953	0.3421
1738	BovineHD2700013197	27	45359212	0.2071
1739	BovineHD2800000092	28	1228577	0.4693
1740	BovineHD2800000602	28	2609804	0.4983
1741	BovineHD2800001104	28	3691153	0.4458
1742	BovineHD2800002014	28	6727906	0.3157
1743	BovineHD2800002758	28	9061846	0.2535
1744	ARS-BFGL-NGS-88529	28	9648302	0.3062
1745	Hapmap41046-BTA-64696	28	10282352	0.2784
1746	ARS-BFGL-NGS-3616	28	10462415	0.4045
1747	BovineHD2800003437	28	11509205	0.3176
1748	Hapmap54617-rs29021286	28	11869873	0.1527
1749	BTA-64815-no-rs	28	13563736	0.1425
1750	BovineHD2800004141	28	14618947	0.4681
1751	ARS-BFGL-NGS-111209	28	17079212	0.1965
1752	BovineHD2800004982	28	17944467	0.4531
1753	ARS-BFGL-NGS-13573	28	18061206	0.1174
1754	ARS-BFGL-NGS-110992	28	20339954	0.3310
1755	Hapmap52240-rs29013844	28	20371221	0.0919
1756	BTB-00980887	28	22495019	0.2338
1757	ARS-BFGL-NGS-11806	28	24394927	0.1508
1758	BovineHD2800013739	28	25291332	0.2802
1759	ARS-BFGL-NGS-75391	28	25878246	0.1607
1760	BovineHD2800007688	28	28734606	0.3249
1761	BovineHD2800007828	28	29308618	0.4006
1762	ARS-BFGL-NGS-8471	28	29381935	0.2491
1763	BovineHD2800007935	28	29734542	0.2300
1764	BovineHD2800008493	28	32054421	0.4026
1765	ARS-BFGL-NGS-76330	28	33248298	0.3243
1766	ARS-USMARC-Parent-EF026086-rs29013660	28	35132199	0.1026
1767	BovineHD2800010126	28	36837461	0.3370
1768	BovineHD2800010973	28	39266326	0.1536
1769	ARS-BFGL-NGS-58930	28	41074012	0.3863
1770	ARS-BFGL-NGS-37392	28	42038658	0.2350
1771	BovineHD2800012123	28	42491873	0.4894

1772	BovineHD2800012862	28	44171402	0.3230
1773	BovineHD2800013264	28	45340536	0.2564
1774	BovineHD2800013311	28	45433793	0.4339
1775	ARS-BFGL-NGS-17624	28	45712332	0.1206
1776	BovineHD2100008217	28	45919071	0.4417
1777	BovineHD2900000116	29	872781	0.2390
1778	BovineHD2900000331	29	1600272	0.4549
1779	BovineHD2900015217	29	2846840	0.1712
1780	BovineHD2900000966	29	3348267	0.4785
1781	BovineHD2900001027	29	3577486	0.2914
1782	BovineHD2900001343	29	4729052	0.4922
1783	BovineHD2900001462	29	5125226	0.3518
1784	BovineHD2900015285	29	9950629	0.3115
1785	BovineHD2900002959	29	10241472	0.4703
1786	ARS-BFGL-NGS-25195	29	11034651	0.2236
1787	BovineHD2900003769	29	12552752	0.4960
1788	BovineHD2900004537	29	15285489	0.4408
1789	BovineHD2900004962	29	16500637	0.4210
1790	BovineHD4100018914	29	20254763	0.4602
1791	BovineHD2900006841	29	23624063	0.4095
1792	BovineHD2900015394	29	24132953	0.3280
1793	BovineHD2900008322	29	27848144	0.3198
1794	ARS-USMARC-Parent-EF034080-rs29024749	29	28278699	0.2361
1795	BovineHD2900009320	29	30722151	0.4981
1796	ARS-BFGL-NGS-109714	29	31764186	0.1142
1797	BovineHD2900009799	29	32285014	0.0683
1798	UA-IFASA-6129	29	34835983	0.1578
1799	BovineHD2900011186	29	36418037	0.4922
1800	ARS-BFGL-NGS-56408	29	37168472	0.1426
1801	BovineHD2900015534	29	41342454	0.4493
1802	BovineHD2900015544	29	42172623	0.4276
1803	BovineHD2900015547	29	42336504	0.2045
1804	BovineHD2900013405	29	44234091	0.4677
1805	BovineHD2900013565	29	44915360	0.4919
1806	BovineHD2900014068	29	46895443	0.4371
1807	BovineHD2900014555	29	48737439	0.3647
1808	BovineHD2900014597	29	48943485	0.4612
1809	BovineHD2900014634	29	49193480	0.4099
1810	BovineHD2900014733	29	51054077	0.4324

Appendix 2. Age-difference based prior probability ratios as a heatmap.



M: maternal parent; P: paternal parent; FS: full siblings; MS: maternal siblings (full + half); PS: paternal siblings.

REFERENCES

Aguilar, I., I. Misztal, S. Tsuruta, A. Legarra, and H. Wang. 2014. PREGSF90–POSTGSF90: Computational tools for the implementation of single-step genomic selection and genome-wide association with ungenotyped individuals in BLUPF90 programs. Pages 1–4 in World Congress on Genetics Applied to Livestock Production (WCGALP). American Society of Animal Science, Vancouver, Canada.

Calus M. P., H. A. Mulder, and J. W. Bastiaansen. 2011. Identification of Mendelian inconsistencies between SNP and pedigree information of sibs. Genet. Sel. Evol. 43:1–12. https://doi.org/10.1186/1297-9686-43-34.

Clarke, S. M., H. M. Henry, K. G. Dodds, T. W. Jowett, T. R. Manley, R. M. Anderson, and J. C. McEwan. 2014. A high throughput single nucleotide polymorphism multiplex assay for parentage assignment in New Zealand sheep. PLoS One. 9:e93392. https://doi.org/10.1371/journal.pone.0093392.

Clayton, D. 2019. snpStats: SnpMatrix and XSnpMatrix classes and methods. R package version 1.36.0. https://doi.org/doi:10.18129/B9.bioc.snpStats.

Dussault, F. M., and E. G. Boulding. 2018. Effect of minor allele frequency on the number of single nucleotide polymorphisms needed for accurate parentage assignment: a methodology illustrated using Atlantic salmon. Aquac. Res. 49:1368–1372. https://doi.org/10.1111/are.13566.

Flanagan, S. P., and A. G. Jones. 2019. The future of parentage analysis: From microsatellites to SNPs and beyond. Mol. Ecol. 28:544–567. https://doi.org/10.1111/mec.14988.

GeneBank. Assembly ARS-UCD1.2. 2018. Available in <https://www.ncbi.nlm.nih.gov/assembly/GCA_002263795.2>. Accessed in May 8, 2021.

Grashei, K. E., J. Ødegård, and T. H. Meuwissen. 2018. Using genomic relationship likelihood for parentage assignment. Genet. Sel. Evol. 50:1–11. https://doi.org/10.1186/s12711-018-0397-7.

Harlizius, B., M. S. Lopes, N. Duijvesteijn, L. H. P. van de Goor, W. A. van Haeringen, H. Panneman, S. E. F. Guimarães, J. W. M. Merks, E. F. Knol. 2011. A single nucleotide polymorphism set for paternal identification to reduce the costs of trait recording in commercial pig breeding, J. Anim. Sci. 89:1661–1668. https://doi.org/10.2527/jas.2010-3347.

Huisman J. 2017. Pedigree reconstruction from SNP data: parentage assignment, sibship clustering and beyond. Mol. Ecol. Resour. 17:1009–1024. R package version 2.0.7. https://doi.org/10.1111/1755-0998.12665.

International Society for Animal Genetics (ISAG). 2012. Cattle Molecular Markers and Parentage Testing Committee. Accessed Feb. 10, 2021. http://www.isag.us/committees.asp?autotry=true&ULnotkn=true.

Knibb, W., P. Whatmore, R. Lamonta, J. Quinn , D. Powell, A. Elizur, T. Anderson, C. Remilton, and N. H. Nguyen. 2014. Can genetic diversity be maintained in long term mass selected populations without pedigree information? — A case study using banana shrimp *Fenneropenaeus merguiensis*. Aquaculture. 428:71–78. https://doi.org/10.1016/j.aquaculture.2014.02.026.

Kopps, A. M., J. Kang, W. B. Sherwin, and P. J. Palsboll. 2015. How Well Do Molecular and Pedigree Relatedness Correspond, in Populations with Diverse Mating Systems, and Various Types and Quantities of Molecular and Demographic Data?. G3: Genes Genomes Genetics 5:1815–1826. https://doi.org/10.1534/g3.115.019323.

Lachance J., and S. A. Tishkoff. 2013. SNP ascertainment bias in population genetic analyses: why it is important, and how to correct it. BioEssays, 35:780–786. https://doi.org/10.1002/bies.201300014.

Legarra, A., D. A. Lourenco, and Z. Vitezica. 2018. Bases for genomic prediction. Paris, França: INRA, Animal Genetics Department.

Marshall, T. C., J. B. K. E. Slate, L. E. B. Kruuk, and J. M. Pemberton. 1998. Statistical confidence for likelihood-based paternity inference in natural populations. Mol. Ecol. 7:639–655. https://doi.org/10.1046/j.1365-294x.1998.00374.x.

Montgomery G.W., M. J. Campbell, P. Dickson, S. Herbert, K. Siemering, K. R. Ewen-White, P. M. Visscher, and N. G. Martin. 2005. Estimation of the rate of SNP genotyping errors from DNA extracted from different tissues. Twin Res. Hum. Genet. 8:346–352. https://doi.org/10.1375/twin.8.4.346.

Panetto, J. C. D. C., M. A. Machado, M. V. G. da Silva, R. S. Barbosa, G. G. dos Santos, R. de MH Leite, and M. G. C. Peixoto. 2017. Parentage assignment using SNP markers, inbreeding and population size for the Brazilian Red Sindhi cattle. Livest. Sci. 204:33–38. https://doi.org/10.1016/j.livsci.2017.08.008.

Powell, J. E., P. M. Visscher, and M. E. Goddard. 2010. Reconciling the Analysis of IBD and IBS in Complex Trait Studies. Nat. Rev. Genet. 11:800–805. http://doi.org/10.1038/nrg2865.

Premachandra, H. K. A., N. H. Nguyen, and W. Knibb. 2019. Effectiveness of SNPs for parentage and sibship assessment in polygamous yellowtail kingfish Seriola lalandi. Aquaculture. 499:24–31. https://doi.org/10.1016/j.aquaculture.2018.09.022.

Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. A. Ferreira, D. Bender, J. Maller, P. Sklar, P. I. W. de Bakker, M. J. Daly, and P. C. Sham. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81:559–575. https://doi.org/10.1086/519795.

Sargolzaei, M., J. P. Chesnais, and F. S. Schenkel. 2014. A new approach for efficient genotype imputation using information from relatives. BMC Genomics. 15, 478. http://dx.doi.org/10.1186/1471-2164-15-478.

Saunders I.W., J. Brohede, and G. N. Hannan. 2007. Estimating genotyping error rates from Mendelian errors in SNP array genotypes and their impact on inference. Genomics, 90:291–296. https://doi.org/10.1016/j.ygeno.2007.05.011.

Speed, D., and D. J. Balding. 2015. Relatedness in the post-genomic era: is it still useful?. Nat. Rev. Genet. 16: 33–44. https://doi.org/10.1038/nrg3821.

Strucken, E.M., B. Gudex, M. H. Ferdosi, H. K. Lee, K. D. Song, J. P. Gibson, M. Kelly, E. K. Piper, L. R. Porto-Neto, S. H. Lee, and C. Gondro. 2014. Performance of different SNP panels for parentage testing in two East Asian cattle breeds. Anim. Genet. 45:572–575. https://doi.org/10.1111/age.12154.

Strucken, E. M., S. H. Lee, H. K. Lee, K. D. Song, J. P. Gibson, and C. Gondro. 2016. How many markers are enough? Factors influencing parentage testing in different livestock populations. J. of Ani. Breed. Genet. 133:13–23. https://doi.org/10.1111/jbg.12179.

VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. J. Dairy Sci. 91:4414–4423. https://doi.org/10.3168/jds.2007-0980.

Weir, B. S., A. D. Anderson, and A. B Hepler. 2006. Genetic relatedness analysis: modern data and new challenges. Nat. Rev. Genet. 7:771–780. https://doi.org/10.1038/nrg1960.

Wiggans G. R., T. S. Sonstegard, P. M. VanRaden, L. K. Matukumalli, R. D. Schnabel, J. F. Taylor, F. S. Schenkel, and C. P. Van Tassell. 2009. Selection of single-nucleotide polymorphisms and quality of genotypes used in genomic evaluation of dairy cattle in the United States and Canada. J. Dairy Sci. 92:3431–3436. https://doi.org/10.3168/jds.2008-1758.

Wiggans, G. R., P. M. VanRaden, L. R. Bacheller, M. E. Tooker, J. L. Hutchison, T. Cooper, and T. S. Sonstegard. 2010. Selection and management of DNA markers for use in genomic evaluation. J. Dairy Sci. 93:2287–2292. https://doi.org/10.3168/jds.2009-2773.

Wright, S. 1922. Coefficients of Inbreeding and Relationship. Am. Nat. 56:330–38. https://doi.org/10.1086/279872.

Chapter 3

LINKAGE DISEQUILIBRIUM, EFFECTIVE POPULATION SIZE AND ROH-BASED INBREEDING COEFFICIENTS IN GIR CATTLE

A. O. Garcia[1], J. C. do C. Panetto[2], P. I. Otto[3], M. V. G. B. da Silva[2], S. E. F. Guimarães[1], and M. A. Machado[2]

[1]Department of Animal Science, Universidade Federal de Viçosa, Viçosa, MG, 36570-900, Brazil

[2]EMBRAPA - Dairy Cattle Research Center, Juiz de Fora, MG, 36038-330, Brazil

[3]Department of Animal Science, Universidade Federal de Santa Maria, Santa Maria, RS, 97105-900, Brazil

ABSTRACT

Besides the productivity, the breeder has to focus also on the genetic makeup of populations, as genetic changes are important either to guide genetic interventions or to evaluate the results of a breeding program. We aim was to calculate the effective size of the genomic population over the generations through linkage disequilibrium, and calculate the inbreeding coefficient based on runs of homozygosity (ROH) to assess the population structure and to investigate the genomic diversity within the Gir breed. In which we detected a decrease in the effective population size, and levels of inbreeding, especially those based on short segments of ROH, with moderate to high values, suggesting the presence of bottlenecks in the genome of Gir cattle. On the other hand, a lower percentage of ROH resided in the longest ROH classes. Breeding strategies that minimize inbreeding and do not make massive use of a few bulls with high genetic value are suggested to maintain genetic variability in future generations. In addition, we strongly recommend reducing the generation interval to maximize genetic progress and increasing the effective population size.

Keywords: genetic diversity, population history, runs of homozygosity

INTRODUCTION

The Gir breed is originally from India and it is currently one of the main cattle breeds used, pure or crossed, for milk production on pasture, especially in tropical regions, due to its low nutritional requirement and resistance to parasites and high temperatures. The first animals of these breed arrived in Brazil between 1870 and 1962, when less than 700 animals were imported (Santana Jr et al., 2014).

The Brazilian Zebu Breeders Association (ABCZ), officially started registering Zebu animals in 1938, encouraging the perpetuation of the Gir breed in the country, initially selected for dual purposes (meat and milk production). In 1985, selection of animals for milk production was intensified; in this context the Brazilian Dairy Gir Breeding Program (PNMGL) and the progeny testing program were created, contributing to the growth of the population of Gir cattle. The first genetic evaluation of milk and fat production was only published in 1993, when a group of only nine Gir bulls was evaluated. From 2001, blood and semen collection began to constitute a DNA bank, aiming at the evaluation of molecular traits in the breed. However, genomic analysis was only implemented in 2016, aiming the genetic evolution of herds.

The dissemination of the Gir breed and the improvement in genetic analysis, the demand for genetically superior sires has increased significantly, making the intensive use of a restricted number of common animals. In addition to this, the introduction of few animals in the country results in an increase of inbreeding and a reduction in the genetic diversity of the breed.

Besides the productivity, the breeder has to focus also on the genetic makeup of populations, as genetic changes are important either to guide genetic interventions or to evaluate the results of a breeding program. The success of breeding programs depends on knowledge of several factors that potentially interfere with the selection, such as effective population size, generation interval and genetic variability (Malhado et al., 2010). Furthermore, knowledge of the population structure combined with information on genetic changes in the population can guide future management actions, allowing the development of strategies that promote genetic gain and the adaptation of a breed to a specific region (Malhado et al., 2009).

The main studies that assess the population structure are carried out with analyzes based on pedigree records (Leroy et al., 2013; Santana Jr et al., 2014, Santana Jr et al., 2016). However, recent genomics approaches allowed the researchers to obtain more information about genetic diversity, population history and inbreeding coefficient (Peripolli et al., 2018; Ospina et al., 2019).

The aim of this study was to calculate genomic effective population size, based on the linkage disequilibrium, throughout generations and calculate the inbreeding coefficient based on runs of homozygosity to investigate the genomic diversity within the Gir breed.

MATERIAL AND METHODS

Data

The animals used in this work are from Gir cattle. All genotyped animals are from a breeding program in Brazil, in which Embrapa Dairy Cattle (Juiz de Fora, Minas Gerais, Brazil) is responsible for genetic evaluations in cooperation with the Brazilian Association of Dairy Gir Breeders (ABCGIL) and the Brazilian Association of Zebu Breeders (ABCZ).

The imputed genotypes were provided by Embrapa Dairy Cattle. The SNP markers used had a call rate > 90%, a minor allele frequency > 0.02 and a P-value for the Hardy-Weinberg balance test > 10^{-6}. In addition, there were no SNPs with the same physical position or unknown. Only autosomal SNP were used. Samples with a call rate < 90% were removed. Imputation was implemented using the FIMPUTE 2.2 software (Sargolzaei, Chesnais & Schenkel, 2014), and lower density panels were imputed to the HD level gradually. The animals genotyped with the Bovine HD BeadChip (Illumina Inc., San Diego, CA, USA) were used as reference population for imputation. Imputation accuracy was 0.9669. Where 21,656 Gir animals and 415,855 SNPs markers remained, which were used in subsequent analyzes.

Linkage Disequilibrium (LD) and LD Decay

The linkage disequilibrium between SNPs was computed as the correlation of gene frequencies (r_{ij}^2) (Hill & Robertson, 1968) using PLINK 1.9 software (Purcell et al., 2007):

$$r_{ij}^2 = \frac{(p_{ij} - p_i q_j)^2}{p_1 p_2 q_1 q_2}$$

where p_{ij} is the probability of the marker allele pair i and j; p_i and q_i are the marginal allelic frequencies at the i and j; $p_1 p_2 q_1 q_2$ is the product of the 4 allele frequencies at both loci.

To assess the decay of the LD, the r_{ij}^2 will be regressed in the distance between the pairs of markers based on the nonlinear parametric model described by Sved (1971):

$$LD_{ij} = \left(\left[1 + 4\beta d_{ij}\right] + e_{ij}\right)^{-1}$$

where LD_{ij} is the r_{ij}^2 observed between SNPs i and j; d_{ij} is the distance in Kb between SNPs i and j; β is the coefficient that describes the LD decay with distance; and e_{ij} is a random residue defined as $e_{ij} \sim N(0, \sigma^2)$.

Effective Population Size

The PLINK 1.9 software (Purcell et al., 2007) was used to calculate the pairwise LD across the chromosomes, based on the correlation coefficient (r_{ij}^2). For all pairs of autosomal SNPs, r_{ij}^2 measures were calculated. The effective population size (N_e) was estimated based on the relationship between N_e, LD without mutation and recombination rate (c) proposed by Sved (1971), which uses the equation (Hayes et al., 2003; Ospsina et al., 2019):

$$N_T = \left({}^{1}/_{4c}\right) \times \left(\left({}^{1}/_{r^2}\right) - 1\right)$$

where *c* represents the distance from the map in Morgans, assuming that 1 cM is equal to 1 Mb; and N_T is the effective population size in the T^{th} generation, and T is ${}^{1}/_{2c}$.

Runs of Homozygosity and Inbreeding Coefficient

Runs of Homozygosity (ROH) were identified in every individual using PLINK 1.9 (Purcell et al., 2007), which uses a sliding window of a specified length or number of homozygous SNPs to scan along each individual's genotype at each SNP marker position to detect homozygous segments. The parameters and thresholds applied to define a ROH were i. a sliding window of 50 SNPs across the genome; ii. the proportion of homozygous overlapping windows was 0.05; iii. the minimum number of consecutive SNPs included in a ROH was 120; iv. the minimum length of a ROH was set to 1 Mb; v. the maximum gap between consecutive homozygous SNPs was 1000 Kb; vi. a density of one SNP per 50 Kb; and vii. a maximum of 1 SNPs with missing genotypes and up to one heterozygous genotype were allowed in a ROH. The ROH were defined by a minimum of 1 Mb in length to avoid short and common ROH that occur throughout the genome due to LD. ROH were classified into five length classes: 1–2, 2–4, 4–8, 8–16, and >16 Mb, identified as $ROH_{1\text{–}2\ Mb}$, $ROH_{2\text{–}4\ Mb}$, $ROH_{4\text{–}8\ Mb}$, $ROH_{8\text{–}16\ Mb}$, and $ROH_{>16\ Mb}$, respectively.

Genomic inbreeding coefficients based on ROH (F_{ROH}) were estimated for each animal, using the *detectRUNS* package (Biscarini et al., 2019) in R software (R version 4.0.2; R Foundation for Statistical Computing, Vienna, Austria), according to McQuillan et al. (2008):

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

where L_{ROHj} is the length of ROH_j, and L_{total} is the total size of each class, chromosome, or wide genome, covered by markers (depending on whether the coverage information corresponds to the class, chromosome or the genome-wide). L_{total}, for the wide genome, was taken to be 2.487.068.108 bp, based on the consensus map.

For each animal F_{ROH} ($F_{ROH1–2\ Mb}$, $F_{ROH2–4\ Mb}$, $F_{ROH4–8\ Mb}$, $F_{ROH8–16\ Mb}$, and $F_{ROH > 16\ Mb}$) was calculated based on ROH distribution of five minimum different lengths (ROH_j): 1–2, 2–4, 4–8, 8–16, and > 16 Mb, respectively. The F_{ROH} for each generation was calculated using the expected length of an autozygous segment in a distribution with mean equal to $1/2g$ Morgans, where g is the number of generations since the common ancestor (Howrigan, Simonson & Keller, 2011). The generation interval, defined as the average age of parents on the birth of their offspring, was also calculated using the *optiSel* (Wellmann, 2021) package in the R software (R version 4.0.2; R Foundation for Statistical Computing, Vienna, Austria).

The inbreeding coefficient was also calculated per chromosome, and the type of distribution was evaluated based on skewness and kurtosis, using the *moments* package (Komsta & Novomestky, 2015) in R software (R version 4.0.2; R Foundation for Statistical Computing, Vienna, Austria).

RESULTS AND DISCUSSION

Linkage disequilibrium and effective population size

We calculate the linkage disequilibrium (LD) based on the r^2, proposed by Hill and Robertson (1968) as the correlation of gene frequencies. The r^2 was used because it is considered a more robust measure of LD than D′, because it is not inflated when estimated from small samples or when there is lower allele frequency (Veroneze et al., 2013). O'Brien et al. (2014) showed that r^2 levels were not affected by the different MAFs, suggesting that the lower LD levels are typical of the indicine breeds.

The graph with the curve of the LD decay pattern in the Gir breed was created based on the estimated r^2 measurements, using the equation described by Sved (1971), and can be seen in Figure 1.

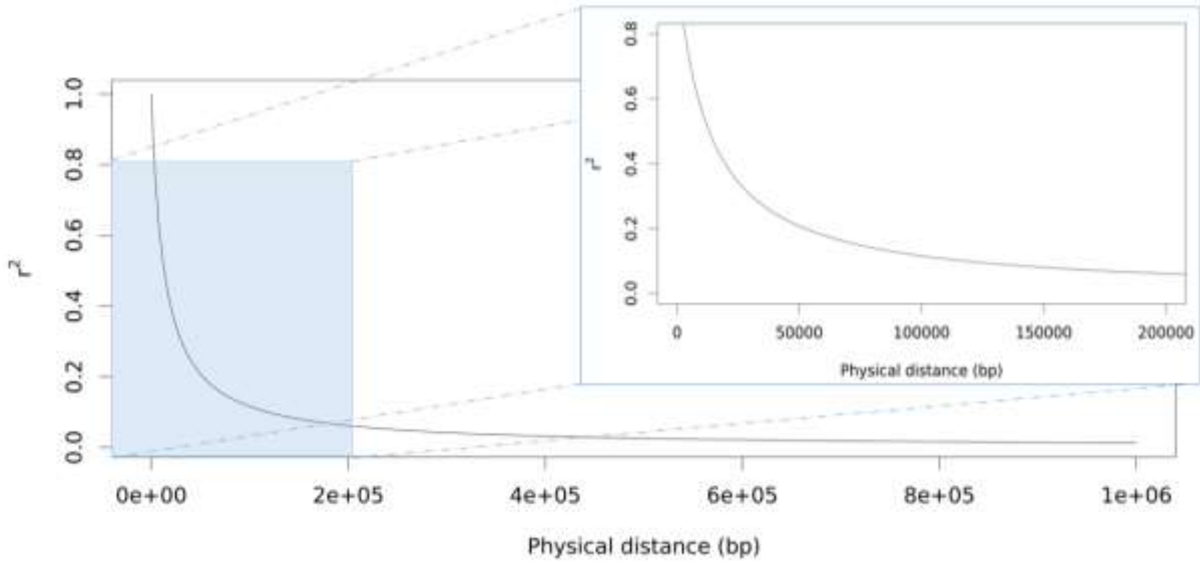


Figure 1. Linkage disequilibrium estimated decay in Gir cattle.

Porto-Neto, Kijas & Reverter (2014) reported the LD decay for taurine, indicine and crossed breeds, in which they observed three distinct curves grouping according to the subspecies. The authors observed that the decay was rapid, with an r^2 below 0.2 for pairs of markers separated by 50 Kb for all breeds, but at distances less than 10 Kb, the taurine breeds had a higher LD ($r^2 = 0.45$) than that indicine them ($r^2 = 0.25$) and crossed ($r^2 = 0.32$). In our work, Gir breed presented an estimated r^2 between pairs of markers equal to 0.2076 for a distance equal to 50 Kb, also showing a rapid decay.

Our average value of r^2 observed at distances less than 10 Kb was equal to 0.3765 ± 0.3538, which is higher than that found by Porto-Neto, Kijas & Reverter (2014) for indicine breeds. This higher LD value can be explained by the smaller effective population size found for our population, which possibly reflects a stronger bottleneck during the formation of the breed.

The mean values of r^2 observed decreased with increasing distance, as well as the percentage of r^2 greater than 0.3 (Table 1). The highest average value of r^2 was equal to 0.4055 ± 0.3630, and the lowest value was equal to 0.0609 ± 0.0926, at distances of 0 – 5 Kb and 900 – 1000 Kb, respectively.

Table 1. Linkage disequilibrium through different inter-marker distances in Gir cattle.

Distance (Kb)	Number of pairs of markers	r^2 mean ± sd*	r^2 median	% r^2 > 0.3
0 – 1	47611	0.3870 ± 0.3566	0.2545	46.13
1 – 5	424702	0.3666 ± 0.3509	0.2275	43.68

5 – 10	458809	0.3506 ± 0.3434	0.2117	41.87
10 – 15	445668	0.3205 ± 0.3311	0.1806	38.27
15 – 20	434283	0.2972 ± 0.3202	0.1586	35.37
20 – 25	427257	0.2802 ± 0.3112	0.1434	33.36
25 – 30	422540	0.2646 ± 0.3025	0.1305	31.39
30 – 40	830067	0.2454 ± 0.2908	0.1161	28.88
40 – 50	817884	0.2249 ± 0.2764	0.1026	26.23
50 – 60	807588	0.2083 ± 0.2645	0.0917	23.94
60 – 70	798003	0.1946 ± 0.2533	0.0836	22.13
70 – 80	793752	0.1819 ± 0.2423	0.0769	20.45
80 – 90	789646	0.1719 ± 0.2337	0.0712	19.06
90 – 100	787392	0.1628 ± 0.2252	0.0667	17.71
100 – 120	1563713	0.1507 ± 0.2136	0.0609	16.00
120 – 140	1554374	0.1379 ± 0.2000	0.0554	14.14
140 – 160	1543415	0.1276 ± 0.1884	0.0511	12.62
160 – 180	1532258	0.1196 ± 0.1789	0.0479	11.48
180 – 200	1522126	0.1122 ± 0.1702	0.0449	10.44
200 – 220	1520179	0.1066 ± 0.1625	0.0429	9.58
220 – 250	2270877	0.1004 ± 0.1543	0.0406	8.65
250 – 275	1886106	0.0949 ± 0.1463	0.0387	7.84
275 – 300	1880176	0.0909 ± 0.1407	0.0374	7.25
300 – 350	3737296	0.0859 ± 0.1327	0.0357	6.49
350 – 400	3719133	0.0806 ± 0.1244	0.0340	5.71
400 – 500	7361193	0.0754 ± 0.1160	0.0322	4.96
500 – 600	7252905	0.0704 ± 0.1076	0.0305	4.27
600 – 700	7008507	0.0666 ± 0.1017	0.0291	3.74
700 – 800	6502810	0.0639 ± 0.0969	0.0280	3.35
800 – 900	5655991	0.0619 ± 0.0940	0.0273	3.10
900 – 1000	4630488	0.0609 ± 0.0926	0.0268	2.99

*sd = standard deviation

O'Brien et al. (2014) evaluated taurine and indicine breeds, including Gir cattle, and observed that lower r^2 levels were found for all indicine breeds compared to taurine breeds, with greater differences in the first 100 Kb. These author suggested the use of high densities of SNP when applying LD-based methods in indicine breeds, compared to most taurines. Within indicine breeds, the levels of r^2 in a distance of 0 – 1 Kb is approximately 0.55, in which the Gir breed presented an average r^2 equal to ≈0.3 with only 16.5 Kb, which would indicate a rapid decay of the LD curve (O'Brien et al., 2014). In current research, the average value of r^2 observed at 0 – 1 Kb was 0.3870 ± 0.3566. To construct the LD decay graph we used the estimated r^2 values (proposed by Sved in 1971), and the distance found for r^2 = 0.3 was 29 Kb, which would be similar to the distance found for taurine breeds found by these O'Brien et al. (2014), that obtained an average distance of 33 Kb.

Considering greater distances (100 – 1000 Kb), the values found by O'Brien et al. (2014) for indicine in a high density data set were similar to ours, with an observed average r^2 of approximately 0.15 (100 Kb) and less than 0.1 for 1Mb, showing smooth decay over greater distances.

When we talk about chromosomes, of the 29 autosome chromosomes analyzed, the chromosome with the lowest average of r^2 was BTA9, with an average of 0.1198 ± 0.1843, followed by BTA5 (0.1162 ± 0.1837). The percentage of r^2 greater than 0.3 in both was equal to 10.97% and 10.48%, respectively. While the autosomes with the lowest averages were BTA25 (0.0859 ± 0.1497) and BTA26 (0.0861 ± 0.1492), with a percentage of r^2 greater than 0.3 equal to 6.48% and 6.49%, respectively. The r^2 values obtained differed among the bovine autosomes and can be seen in Appendix 1.

Espigolan et al. (2013) reported that the mean r^2 in autosome chromosomes in Nellore cattle ranged from 0.003 ± 0.01 (BTA29) to 0.21 ± 0.26 (BTA13). While Lu et al. (2012) reported values ranging from 0.26 ± 0.29 (BTA14) to 0.18 ± 0.22 (BTA29) for Angus cattle. When compared with the values found in the two studies, our results, regarding the Gir breed, were more similar to the results found for Nellore cattle. This result was expected, since animals of the indicine breeds have shown lower values of r^2 than those of the taurine breeds (McKay et al., 2007; O'Brien et al., 2014; Porto-Neto et al., 2014). Higher LD values in taurine breeds are generally justified by a smaller effective population size and greater genetic bottleneck exercised under these animals in remote times.

Differences in LD levels can be attributed to historical population events, which include mixing with wild-type ancestors (Murray et al., 2010), less selection pressure (Thévenon et al., 2007) and larger effective population sizes (Ospina et al., 2019; Santana Jr et al., 2014) in the indicine breeds.

According to Falconer and Mackay (1996), the effective population size (N_e), or the effective number of breeders, represents the number of individuals that give rise to the calculated sample variance, or inbreeding rate, if they reproduce in the way of the population idealized. The N_e, based on the genomic data, in the corresponding generation, can be estimated based on the LD standards (Hayes et al., 2003; Barbato et al., 2015; Jasielczuk, 2020). The strength of LD at different genetic distances between loci allows us to infer about an effective ancestral population size. The N_e estimated for the Gir cattle population for the 50 previous generations is illustrated in Figure 2.

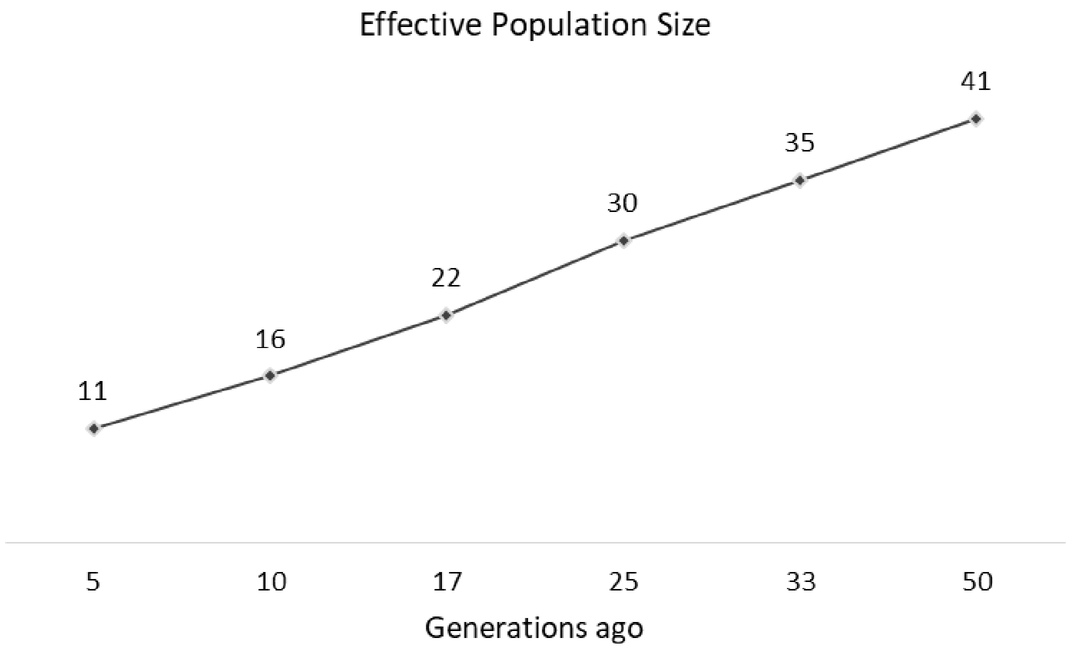


Figure 2. Estimated effective size of the population in Gir cattle in the last fifty generations.

The estimated N_e was higher in generation 50, with 41 animals, and declined drastically over the years; the last generation (5) presented eleven animals. The N_e estimates obtained for the Gir cattle exhibit a decreasing trend over the generations (analyzing from the oldest to the most recent generation) probably due to the loss of genetic variability in the population. According to FAO (1992), the rate of genetic diversity loss increases drastically at an effective population size smaller than one hundred.

Malhado et al. (2010) reported N_e, based on the sex ratio, for Gir cattle in Northeastern Brazil estimated on 100 to 150 at period from 2002 to 2006, corroborating with Reis-Filho et al. (2010) and Santana Jr et al. (2014), who found similar values of N_e, based on inbreeding coefficients. Our population of Gir dairy cattle had an effective size much smaller than that found by other authors in the same breed. The difference was possibly due to the methodology used in the analysis since unlike the others, which used information provided by the pedigree, we used information from molecular markers, resulting in values that are likely to be more accurate.

The average generation interval was 9.08 years. Higher value for Gir animals than those found by Malhado et al. (2010) and Reis-Filho et al. (2010), which estimated 7.9 and 8.41 years, respectively. Longer generation intervals result in reduced annual genetic gain and, consequently, lower economic returns for the breeding program. Generally, the long generation intervals are partly due to the longevity of the Gir breed, with frequent cases of animals that are kept for reproduction until old age, in addition to the continuous use of specific bulls, without replacement. With the implementation of genomics approach in breeding programs, much younger animals have been included in the tests, which is a great incentive for improvements, in addition to the fast replacement of the bull and decreasing the age at the first calving of the dams.

Runs of homozygosity and inbreeding coefficient

A total of 1,156,117 ROHs were found among the 21,656 Gir cattle analyzed in this study. An average of 53 ± 9.08 ROHs were found per animal, with a maximum number equal to 284 and minimum of 6 ROHs. The average length of ROHs per individual was equal to 3,237.37 Kb ± 1,056.99 Kb. The majority of ROHs were between 1 and 2 Mb length, with 61.39%, present at least once in all animals (Table 2). The proportion of ROHs longer than 16 Mb was 2.29%, with an average length of 25.68 Mb. A total of 12,272 animals exhibited ROHs in the >16 Mb length class, with an average of 2 ROHs per animal in this class and a minimum and maximum equal to 1 and 19 ROHs, respectively.

Table 2. Descriptive statistics of runs of homozygosity (ROH) number and length (in Mb) by ROH length class ($ROH_{1-2\,Mb}$, $ROH_{2-4\,Mb}$, $ROH_{4-8\,Mb}$, $ROH_{8-16\,Mb}$, and $ROH_{>16\,Mb}$).

Class	N_{ROH}*	Percent (%)	L_{ROH} (Mb)**	S_{ROH}***	Genome coverage
$ROH_{1-2\,Mb}$	709782	61.39	1.1838	32.78	0.05%
$ROH_{2-4\,Mb}$	246853	21.35	2.7673	11.40	0.11%
$ROH_{4-8\,Mb}$	116471	10.07	5.5340	5.38	0.22%
$ROH_{8-16\,Mb}$	56489	4.89	11.1398	2.82	0.45%
$ROH_{>16\,Mb}$	26522	2.29	25.6808	2.16	1.03%

*N_{ROH} = runs of homozygosity number; **L_{ROH} = average length of ROH; ***S_{ROH} = average ROH number per animal

The highest percentage of ROH (61.39%) occurred in the shortest distance ($ROH_{1-2\,Mb}$), which may be mainly due to the small size of the breed population and the strong founding effect due to genetic bottlenecks. In contrast, the lowest percentage of ROH (2.29%) resided in the longest class ($ROH_{>16\,Mb}$), which suggests a rather small reduction in genetic variability that has occurred in the last generations. We can see in the results a marked reduction in the number of ROH in relation to their length, which, consequently, reflects in the reduction of the F_{ROH} over the generations, which is coherent, due to the mating strategies applied over the years to reduce the mating between highly related animals.

The number of ROHs and length of ROH segments varied across chromosomes, as shown in Figure 3. The highest number of ROHs was identified on BTA5 (69,806), followed by BTA1 (67,040), and BTA2 (65,640), while the lowest was identified on BTA28, with 14,815 ROHs detected, BTA27 (15,415), and BTA25 (16,899).

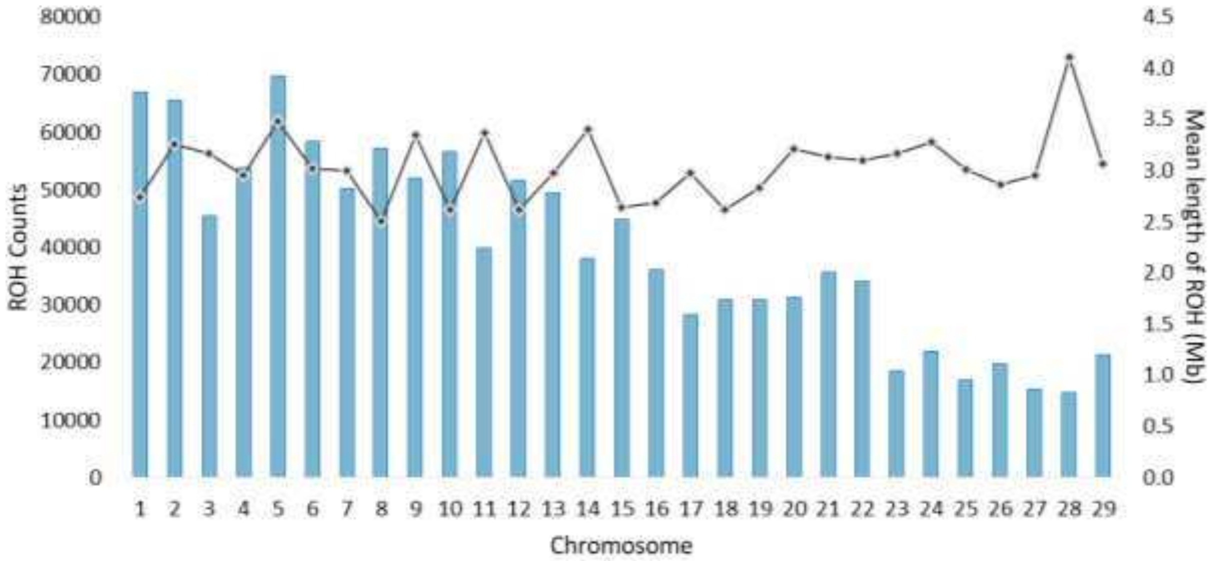


Figure 3. Distribution and average length of runs of homozygosity (ROH) in Mb detected across the autosomal genome in Gir cattle. The bar plots show the ROH counts per chromosome and the gray line shows the average ROH size (Mb) per chromosome.

However, when we identified the coverage by ROHs considering the length of the chromosome (Figure 4), the chromosome with the highest coverage by ROHs was detected in BTA28, with 8.94% coverage by ROH, followed by BTA25 (7.11%), BTA27 (6.48%), BTA23 (6.03 %) and BTA29 (6.00%). The chromosomes with the lowest coverage were BTA1 (1.73%), BTA8 (2.21%), and BTA2 (2.39%).

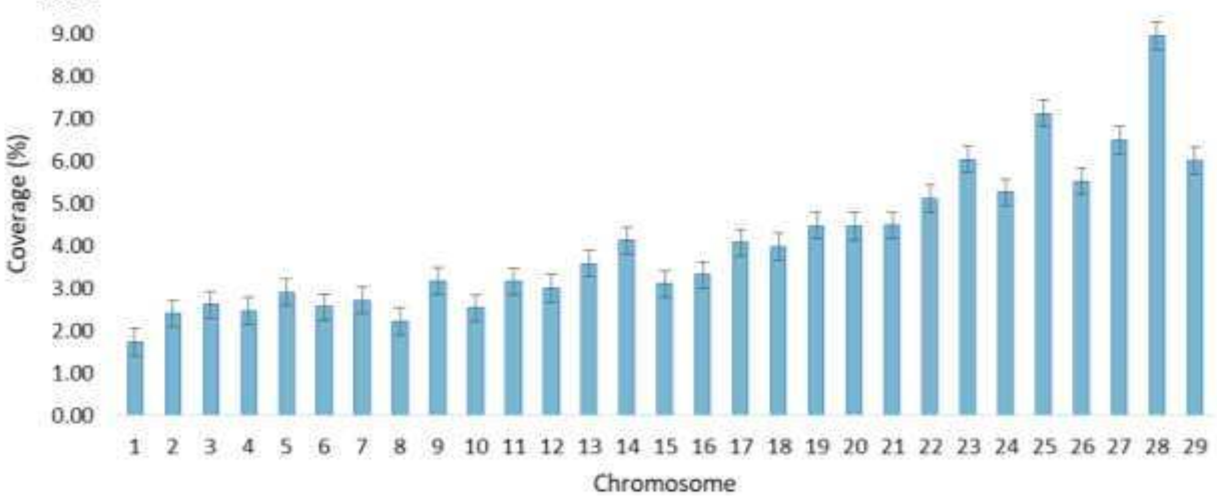


Figure 4. Average percentage of chromosome coverage by runs of homozygosity. The error bars indicate the standard error.

The extent and frequency of ROH have been widely used to infer ancestry at the individual and breed level. The inbreeding, measured as the proportion of the genome covered by ROH, resulted in an average F_{ROH} of 0.0646 ± 0.0287 for our Gir cattle population, the distribution can be seen in Figure 5.

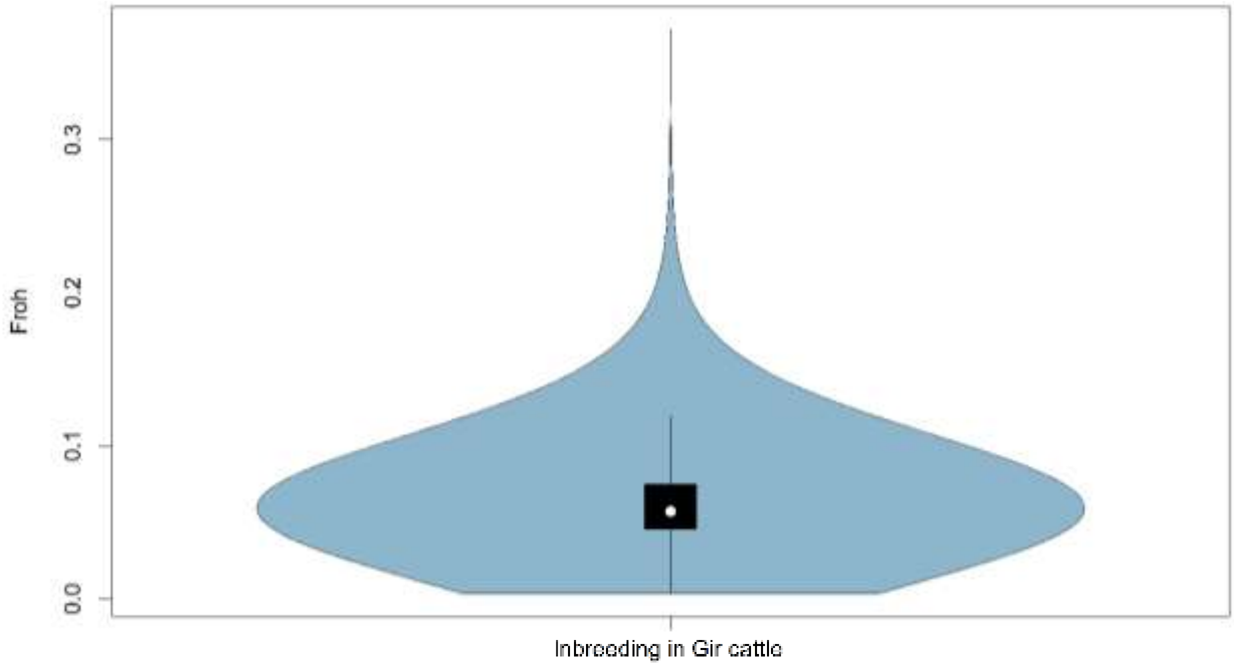


Figure 5. Inbreeding coefficient based on runs of homozygosity (F_{roh}) in Gir cattle.

Santana Jr et al. (2014) and Reis-Filho et al. (2010), analyzing the coefficient of inbreeding based on the pedigree (F_{PED}), also in the Gir breed, found values of 0.0192 and 0.0280, respectively. These studies described inbreeding values close to those found by us at distances above 8 Mb ($F_{ROH8–16\ Mb}$ = 0.0263 ± 0.0250, and $F_{ROH>16\ Mb}$ = 0.0223 ± 0.0211). This was explained by Zavarez et al. (2015), who observed that the incomplete pedigree fails to capture remote inbreeding and estimates based on pedigree information are only comparable with F_{ROH} calculated on large ROH. The variation between these two estimates can be attributed to the fact that the F_{PED} does not assume selection and recombination events in the entire genome, therefore, it does not consider the potential bias of these events (Peripolli et al., 2018). Furthermore, we emphasize that pedigree relatedness is estimated from statistical expectations of the probable genomic proportion of IBD, while genotype-based estimates show the real relationship between individuals and can provide greater accuracy in the relationship (Visscher et al., 2006).

Long ROH are generally considered to be a sign of recent inbreeding, whereas short ROH can capture ancient inbreeding which derived from older ancestors and can capture population bottlenecks. Under the assumption that 1 cM equal to approximately 1 Mb, ROH can be separated in length classes to express different points in time when inbreeding occurred (Curik, Ferenčaković & Sölkner, 2014). To differentiate ancient and recent inbreeding, we calculated F_{ROH} based upon five different ROH length classes, as presented in Table 3. We consider F_{ROH} of 1 Mb in length equal to ≈50 generations, 2 Mb ≈25, 4 Mb ≈12, 8 Mb ≈6, 16 Mb ≈3, and >16 Mb is equivalent to less than 3 generations.

Inbreeding events occurring over ≈35 generations ago were presented by F_{ROH} calculated per length below or equal to 2 Mb. The degree of inbreeding based on the longest ROH class showed an average inbreeding equal to 0.0223 ± 0.0211, reflecting inbreeding events which happened < 3 generations ago.

Table 3. Descriptive statistics of the genomic inbreeding coefficients based on runs of homozygosity (F_{ROH}) for different lengths ($F_{ROH1–2\ Mb}$, $F_{ROH2–4\ Mb}$, $F_{ROH4–8\ Mb}$, $F_{ROH8–16\ Mb}$, and $F_{ROH>16\ Mb}$) for genotyped animals (*n*).

Class	*n*	Mean ± sd*	Median	Minimum	Maximum	CV**
$F_{ROH1–2\ Mb}$	21656	0.0646 ± 0.0287	0.0572	0.0033	0.3722	44.43%
$F_{ROH2–4\ Mb}$	21655	0.0489 ± 0.0285	0.0414	0.0010	0.3606	58.23%
$F_{ROH4–8\ Mb}$	21630	0.0363 ± 0.0277	0.0285	0.0016	0.3483	76.15%
$F_{ROH8–16\ Mb}$	20014	0.0263 ± 0.0250	0.0185	0.0032	0.3338	94.91%
$F_{ROH>16\ Mb}$	12272	0.0223 ± 0.0211	0.0153	0.0064	0.2811	94.56%

*sd = standard deviation; **CV = coefficient of variation

The high values of the coefficient of variation (CV) found here (Table 3) show that there is a big difference between the mean and median for each length of F_{ROH}. Therefore, we can assume that the median, instead of the mean, is the measure of central tendency most suitable for the F_{ROH} coefficients.

The inbreeding calculated per chromosome showed large variation among chromosomes both in terms of average values and the distribution of inbreeding level per individual (Figure 6). The highest average inbreeding was found in BTA24 (0.201 ± 0.205) and the lowest on BTA2 (0.098 ± 0.106). For all chromosomes a highly skewed distribution towards higher values was found, when it was found a distribution skewness above ten. A kurtosis value above three can be used to investigate the presence of outlier individuals per each chromosome. In total, 18 out of the 29 chromosomes showed a kurtosis higher than three, highlighting the presence of some animals exhibiting an excess of homozygosity.

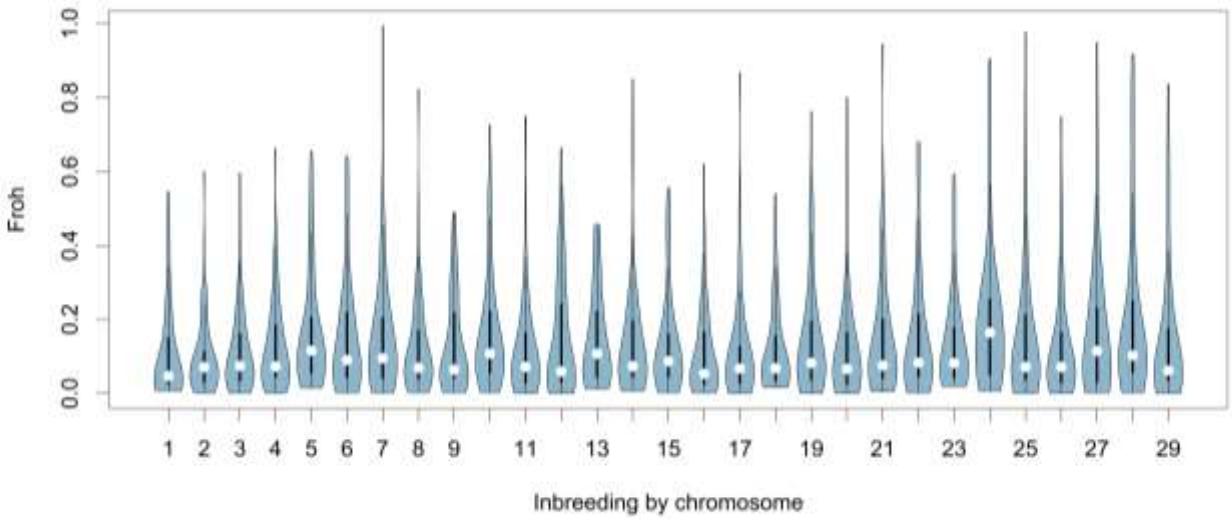


Figure 6. Inbreeding coefficient based on runs of homozygosity (F_{roh}) by chromosome in Gir cattle.

In summary, the linkage disequilibrium, which is the non-random association between the loci, allows the inheritance of complete DNA segments in regions with a low recombination rate. When an animal has homozygous regions, these segments are called runs of homozygosity. The extent and frequency of ROH may provide information about the ancestry of an individual and its population. The frequency of ROH can signal islands, used to detect selection signatures, while the extension can provide information about inbreeding. In which short ROHs indicate old inbreeding and long ROHs indicate recent inbreeding, suggesting about the effective population size.

The Gir population used in this study had the lowest N_e value in the estimated generation of five, suggesting a recent inbreeding that is supported by the occurrence of long ROH found.

CONCLUSION

The population structure of the Gir breed in this study was analyzed via N_e, on LD and ROHs detected. We identified a decrease in the LD-based N_e, when calculated over the last 50 generations, suggesting a recent inbreeding that is supported by the occurrence of long ROH found. The average generation interval in this population was 9.08 years. The ROH analysis provided additional information to further optimize the breeding and management decisions needed to ensure the long term survival of the breed. The average inbreeding based on ROH in Gir cattle was equal to 0.0646 ± 0.0287. Inbreeding levels, especially those based on short ROH segments, were moderate to high, suggesting a small N_e and the presence of past bottlenecks in the genome of Gir cattle. Conversely, the lowest percentage of ROH resided in the longest ROH classes, showing a change in genetic variability during the last generations. Therefore, we recommend the continuous monitoring of conservation measures to minimize the loss of genetic diversity in the breed. Breeding strategies that minimize inbreeding and do not make massive use of a few bulls with high genetic value are suggested to maintain genetic variability in future generations. In addition, we strongly recommend reducing the generation interval to maximize genetic progress and increasing the effective population size.

ACKNOWLEDGMENTS

We thank EMBRAPA Agricultural Informatics Research Center, the Brazilian Association of Dairy Gir Breeders (ABCGIL), Brazilian Association of Zebu Breeders (ABCZ), and the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Brasilia), and Ministério da Ciência Tecnologia e Inovação (MCTI)/Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)/Instituto Nacional de Ciência e Tecnologia–Ciência Animal (INCT-CA, Brasilia, Brazil) for financial support for their financial and technical support. The authors have not stated any conflicts of interest.

APPENDIX

Appendix 1. Linkage disequilibrium across autosomal chromosomes in Gir cattle.

Chr	Pairs of markers (N)	r^2 mean ± sd*	r^2 median	% r^2 > 0.3
BTA1	4486591	0.0958 ± 0.1668	0.0342	8.11
BTA2	3622290	0.1084 ± 0.1739	0.0418	9.54
BTA3	3347797	0.0932 ± 0.1594	0.0340	7.57
BTA4	3148438	0.1019 ± 0.1699	0.0378	8.60
BTA5	2844561	0.1162 ± 0.1837	0.0449	10.48
BTA6	4027514	0.1059 ± 0.1715	0.0398	9.36
BTA7	3276831	0.1031 ± 0.1741	0.0370	8.90
BTA8	3664491	0.0990 ± 0.1632	0.0384	8.13
BTA9	3386160	0.1198 ± 0.1843	0.0479	10.97

BTA10	2631627	0.0964 ± 0.1658	0.0341	8.07
BTA11	2791279	0.0952 ± 0.1638	0.0344	7.79
BTA12	2388984	0.0983 ± 0.1694	0.0350	8.07
BTA13	2116548	0.1074 ± 0.1715	0.0424	9.25
BTA14	2616951	0.1016 ± 0.1635	0.0396	8.58
BTA15	2138630	0.0986 ± 0.1652	0.0363	8.39
BTA16	2625581	0.0975 ± 0.1616	0.0374	8.00
BTA17	2066723	0.0930 ± 0.1603	0.0335	7.53
BTA18	1897406	0.0993 ± 0.1785	0.0340	8.07
BTA19	1528297	0.0987 ± 0.1650	0.0364	8.29
BTA20	1947612	0.1006 ± 0.1665	0.0374	8.60
BTA21	1934475	0.1065 ± 0.1743	0.0387	9.38
BTA22	1578957	0.1001 ± 0.1662	0.0363	8.56
BTA23	1457332	0.0962 ± 0.1620	0.0358	7.90
BTA24	1821411	0.0923 ± 0.1613	0.0327	7.50
BTA25	1109031	0.0859 ± 0.1497	0.0321	6.48
BTA26	1400942	0.0861 ± 0.1492	0.0324	6.49
BTA27	1237917	0.0878 ± 0.1495	0.0334	6.68
BTA28	1204308	0.0928 ± 0.1524	0.0366	7.35
BTA29	1166940	0.1034 ± 0.1710	0.0373	9.13

*sd = standard deviation

REFERENCES

Barbato, M., P. Orozco-terWengel, M. Tapio, and M. W. Bruford. 2015. SNeP: a tool to estimate trends in recent effective population size trajectories using genome-wide SNP data. Front. genet. 6, 109. https://doi.org/10.3389/fgene.2015.00109.

Biscarini F., P. Cozzi, G. Gaspa, and G. Marras. 2019. *detectRUNS*: Detect Runs of Homozygosity and Runs of Heterozygosity in Diploid Genomes. R package version 0.9.6. https://CRAN.R-project.org/package=detectRUNS.

Clayton, D. 2019. *snpStats*: SnpMatrix and XSnpMatrix classes and methods. R package version 1.36.0. https://doi.org/doi:10.18129/B9.bioc.snpStats.

Curik, I., M. Ferenčaković, and J. Sölkner. 2014. Inbreeding and runs of homozygosity: a possible solution to an old problem. Livest. Sci. 166:26–34. https://doi.org/10.1016/j.livsci.2014.05.034.

Espigolan, R., F. Baldi, A. A. Boligon, F. R. Souza, D. G. Gordo, R. L. Tonussi, D. F. Cardoso, H. N. Oliveira, H. Tonhati, M. Sargolzaei, F. S. Schenkel, R. Carvalheiro, J. A. Ferro, and L.

G. Albuquerque. 2013. Study of whole genome linkage disequilibrium in Nellore cattle. BMC Genomics, 14:1–8. https://doi.org/10.1186/1471-2164-14-305.

Falconer, D. S., and T. F. C. Mackay. 1996. Small Populations: II Less simplified conditions. Pages 70 – 76 in Introduction to quantitative genetics, 4th edition. Longman Scientific and Technical, Essex, England.

FAO, 1992. Monitoring animal genetic resources and criteria for prioritization of breeds, K. Maijala. In: Hodges, J. The management of global animal genetic resources, Proceedings of an FAO Expert Consultation, Rome, Nº 104.

GeneBank. Assembly ARS-UCD1.2. 2018. Available in <https://www.ncbi.nlm.nih.gov/assembly/GCA_002263795.2>. Accessed in May 8, 2021.

Hayes, B. J., P. M. Visscher, H. C. McPartlan, and M. E. Goddard. 2003. Novel multilocus measure of linkage disequilibrium to estimate past effective population size. Genome Res. 13:635–643. http://www.genome.org/cgi/doi/10.1101/gr.387103.

Hill, W. G., and A. Robertson. 1968. Linkage disequilibrium in finite populations. Theor. Appl. Genet. 38:226–231. https://doi.org/10.1007/BF01245622.

Howrigan, D. P., M. A. Simonson, and M. C. Keller. 2011. Detecting autozygosity through runs of homozygosity: a comparison of three autozygosity detection algorithms. BMC genomics, 12:1-15. https://doi.org/10.1186/1471-2164-12-460.

Jasielczuk, I., A. Gurgul, T. Szmatoła, E. Semik-Gurgul, K. Pawlina-Tyszko, M. Szyndler-Nędza, T. Blicharski, K. Szulc, E. Skrzypczak, and M. Bugno-Poniewierska. 2020. Comparison of linkage disequilibrium, effective population size and haplotype blocks in Polish Landrace and Polish native pig populations. Livest. Sci. 231, 103887. https://doi.org/10.1016/j.livsci.2019.103887.

Komsta L., and F. Novomestky. 2015. *moments*: Moments, cumulants, skewness, kurtosis and related tests. R package version 0.14. https://CRAN.R-project.org/package=moments.

Leroy, G., T. Mary-Huard, E. Verrier, S. Danvy, E. Charvolin, and C. Danchin-Burge. 2013. Methods to estimate effective population size using pedigree data: Examples in dog, sheep, cattle and horse. Genetics Selection Evolution, 45:1–10. https://doi.org/10.1186/1297-9686-45-1.

Lu, D., M. Sargolzaei, M. Kelly, C. Li, G. Vander Voort, Z. Wang, G. Plastow, S. Moore, and S. Miller. 2012. Linkage disequilibrium in Angus, Charolais, and Crossbred beef cattle. Front. Genet. 3:152. https://doi.org/10.3389/fgene.2012.00152.

Malhado, C. H. M., P. L. S. Carneiro, A. C. M. Malhado, J. A. M. Martins, R. Martins Filho, and R. Bozzi. 2010. History of registered Gyr breed in Brazilian Northeast: population structure and genetic improvement of growth traits. Cienc. Rural, 40:1385–1391. https://doi.org/10.1590/S0103-84782010000600023.

Malhado, C. H. M., P. L. S. Carneiro, R. Martins Filho, and D. M. M. R. Azevedo. 2009. Genetic and populational background of Pure Nelore cattle breed in Brazilian Northeastern Sertão. Pesqui. Agropecu. Bras. 44:713–718. https://doi.org/10.1590/S0100-204X2009000700010.

McKay, S. D., R. D. Schnabel, B. M. Murdoch, L. K. Matukumalli, J. Aerts, W. Coppieters, D. Crews, E. Dias-Neto, C. A. Gill, C. Gao, H. Mannen, P. Stothard, Z. Wang, C. P. Van Tassell, J. L. Williams, J. F. Taylor, and S. S. Moore. 2007. Whole genome linkage disequilibrium maps in cattle. BMC Genet. 8:74. https://doi.org/10.1186/1471-2156-8-74.

McQuillan, R., A. L. Leutenegger, R. Abdel-Rahman, C. S. Franklin, M. Pericic, L. Barac-Lauc, N. Smolej-Narancic, B. Janicijevic, O. Polasek, A. Tenesa, A. K. MacLeod, S. M Farrington, P. Rudan, C. Hayward, V. Vitart, I. Rudan, S. H. Wild, M. G. Dunlop, A. F. Wright, H. Campbell, and J. F. Wilson. 2008. Runs of homozygosity in European populations. Am. J. Hum. Genet. 83:359–372. https://doi.org/10.1016/j.ajhg.2008.08.007.

Murray, C., E. Huerta-Sanchez, F. Casey, and D. G. Bradley. 2010. Cattle demographic history modelled from autosomal sequence variation. Philosophical Transactions of the Royal Society B: Biological Sciences, 365:2531–2539. https://doi.org/10.1098/rstb.2010.0103.

O'Brien, A. M. P., G. Mészáros, Y. T. Utsunomiya, T. S. Sonstegard, J. F. Garcia, C. P. Van Tassell, R. Carvalheiro, M. V. B. da Silva, and J. Sölkner. 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. https://doi.org/10.1016/j.livsci.2014.05.007.

Ospina A. M. T., A. M. Maiorano, R. A. Curi, G. L. Pereira, M. E. Zerlotti-Mercadante, J. N. dos S. Gonçalves Cyrillo, R. R. Aspilcueta-Borquis, and J. A. I. de V. Silva. 2019. Linkage disequilibrium and effective population size in Gir cattle selected for yearling weight. Reprod. in Domest. Anim., 54:1524–1531. https://doi.org/10.1111/rda.13559.

Peripolli, E., N. B. Stafuzza, D. P. Munari, A. L. F. Lima, R. Irgang, M. A. Machado, M. A., J. C. do C, Panetto, R. V. Ventura, F. Baldi, and M. V. G. B da Silva. 2018. Assessment of runs of homozygosity islands and estimates of genomic inbreeding in Gyr (*Bos indicus*) dairy cattle. BMC Genomics, 19:1–13. https://doi.org/10.1186/s12864-017-4365-3.

Porto-Neto, L. R., J. W. Kijas, and A. Reverter. 2014. The extent of linkage disequilibrium in beef cattle breeds using high-density SNP genotypes. Genet. Sel. Evol. 46:1–5. https://doi.org/10.1186/1297-9686-46-22.

Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. A. Ferreira, D. Bender, J. Maller, P. Sklar, P. I. W. de Bakker, M. J. Daly, and P. C. Sham. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81:559–575. https://doi.org/10.1086/519795.

Reis-Filho, J. C., P. S. Lopes, R. D. S. Verneque, R. D. A. Torres, R. L. Teodoro, and P. L. S. Carneiro. 2010. Population structure of Brazilian Gyr dairy cattle. Rev. Bras. Zootec. 39:2640–2645. https://doi.org/10.1590/S1516-35982010001200012.

Santana Jr, M. L., R. J. Pereira, A. B. Bignardi, D. R. Ayres, H. Tonhati, G. D. O. Menezes, L. O. C. Silva, G. Leroy, C. H. C. Machado, L. A. Josahkian, and L. G. Albuquerque. 2016. Structure and genetic diversity of Brazilian Zebu cattle breeds assessed by pedigree analysis. Livestock Science, 187:6–15. https://doi.org/10.1016/j.livsci.2016.02.002.

Santana Jr, M. L., R. J. Pereira, A. B. Bignardi, L. El Faro, H. Tonhati, and L. G. Albuquerque. 2014. History, structure, and genetic diversity of Brazilian Gir cattle. Livestock Science, 163:26–33. https://doi.org/10.1016/j.livsci.2014.02.007.

Sargolzaei, M., J. P. Chesnais, and F. S. Schenkel. 2014. A new approach for efficient genotype imputation using information from relatives. BMC Genomics. 15, 478. http://dx.doi.org/10.1186/1471-2164-15-478.

Sved, J. A. 1971. Linkage disequilibrium and homozygosity of chromosome segments in finite populations. Theor. Popul. Biol. 2:125–141. https://doi.org/10.1016/0040-5809(71)90011-6.

Thévenon, S., G. K. Dayo, S. Sylla, I. Sidibé, D. Berthier, H. Legros, D. Boichard, A. Eggen, and M. Gautier. 2007. The extent of linkage disequilibrium in a large cattle population of western Africa and its consequences for association studies. Anim. Genet. 38:277–286. https://doi.org/10.1111/j.1365-2052.2007.01601.x.

Veroneze, R., P.S. Lopes, S. E. F. Guimarães, F. F. Silva, M. S. Lopes, B. Harlizius, and E. F. Knol. 2013. Linkage disequilibrium and haplotype block structure in six commercial pig lines. J. Anim. Sci. 91:3493–3501. https://doi.org/10.2527/jas.2012-6052.

Visscher, P. M., S. E. Medland, M. A. R. Ferreira, K. I. Morley, G. Zhu, B. K. Cornes, Montgomery G. W., Martin N. G. 2006. Assumption-free estimation of heritability from genome-wide identity-bydescent sharing between full siblings. PLoS Genet. 2:0316–25. https://doi.org/10.1371/journal.pgen.0020041.

Wellmann, R. 2021. optiSel: Optimum Contribution Selection and Population Genetics. R package version 2.0.5. https://CRAN.R-project.org/package=optiSel.

Zavarez, L. B., Y. T. Utsunomiya, A. S. Carmo, H. H. Neves, R. Carvalheiro, M. Ferenčaković, A. M. P. O'Brien, I. Curik, J. B. Cole, C. P. Van Tassell, M. V. G. B. da Silva, T. S. Sonstegard, J. Sölkner, and J. F. Garcia1. 2015. Assessment of autozygosity in Nellore cows (*Bos indicus*) through high-density SNP genotypes. Front. Genet. 6:5. https://doi.org/10.3389/fgene.2015.00005.

Chapter 4

GENERAL DISCUSSION

Introduction

Considering that the estimation of genetic values and studies of the structure and genetic diversity of animal populations have been carried out using traditional methods, involving pedigree records (Patry & Ducrocq, 2011; Santana Jr et al., 2014) in which pedigree information is used to infer on the relationship coefficients and inbreeding of animals, and thereby obtain valuable information on their productive potential, history of population and genetic diversity.

Incorrect or incomplete pedigree information can considerably reduce genetic gain. According to Falconer & Mackay (1996), when we use the traditional relationship matrix, there is an under or overestimation of relationship that will directly interfere in the estimation of the animal's genetic value. However, it is also shown in our work that the use of molecular markers allows the relationship between animals to be investigated and a more precise relationship matrix to be estimated in the absence of a pedigree.

Analyzes of population structure and genetic diversity are traditionally also based on information from the pedigree. Generally, information related to population inbreeding, or even information on the sex ratio of animals, in used to calculate the effective size (N_e) of the population (Reis-Filho et al., 2010; Malhado et al., 2010) and to make inferences about the history and genetic variability of the population in question. Currently, it is possible to estimate these population parameters based on genotype information, where inbreeding and N_e can be obtained based on linkage disequilibrium and runs of homozygosity, which makes the values much closer to the real ones.

In this chapter, I discuss the importance of accurate pedigree information, in addition to the impact of pedigree errors on genetic analysis, and the results and application of genomic analyzes in studies of Gir cattle population structure.

The importance of relationship and parentage analysis in breeding programs

The genetic value, estimated for the selection of animals aiming to genetic progress, is based on information between relatives, through the Best Linear Unbiased Prediction (BLUP), via mixed model equations proposed by Henderson (1975). BLUP uses the traditional relationship matrix (A) between animals, which presents the genetic covariates between individuals (Lopes, Torres & Pires, 2005; Legarra, Lourenco & Vitezica, 2018). However, BLUP is often not used with maximum efficiency due to errors in the pedigree annotation and also due to the difficulty in estimating the real relationship between individuals. According to Falconer & Mackay (1996), when we use the traditional relationship matrix, there is an under or overestimation of relationship that will directly interfere in the estimation of the animal's genetic value. Currently, with the genomic era, this genetic value is estimated via Genomic BLUP (GBLUP), in which it uses the genomic relationship matrix (G). Moreover, recent approaches have been created so that it is possible to combine the information of relatives of genotyped animals and not genotyped, with the Single Step GBLUP (ssGBLUP), where the combined matrix A and G (H matrix) is used.

The relevance of the parentage test is due to losses in the accuracy of selection and reduction of genetic progress due to mistaken records about the paternity of animals (Van Vleck, 1970; Geldermann, Pieper & Weber, 1986). According to Ron et al. (1996), parentage errors can reach up to 20% of the records, which considerably reduces genetic progress in the population in question. This is because in animal breeding, to achieve greater accuracy in the estimation of genetic values and prediction of genetic gain in a herd, a complete and reliable pedigree is necessary, while errors in the pedigree can reduce them (Geldermann, Pieper & Weber, 1986; Senneke et al., 2004). In addition, information about genealogy is indispensable for the study of low-incidence characteristics, such as diseases, genetic defects and mortality (Harlizius, 2011).

In chapter one, it was discussed that incorrect or incomplete pedigree information can reduce genetic gain. However, it has also been shown in our work that using molecular markers, such as Single Nucleotide Polymorphism (SNP), the relationship between animals can be investigated and a more accurate relationship matrix can be estimated in the absence of a pedigree.

The parentage test described in chapter two was performed using the likelihood ratio approach with SNP markers common to four different commercial chips, and the relationship coefficients were obtained based on the imputed genotypes. The main motivations of our work was to ensure the accuracy and depth of the pedigree, and to obtain the genomic relationship between the animals, which present values closer to the real ones, when compared with the traditional relationship.

The likelihood ratio approach to pedigree reconstruction

In this dissertation we use the likelihood ratio (LR) approach to reconstruct the pedigree of Gir cattle, in order to correct any errors in annotation and deepen the information contained in the pedigree records. The LR approach is a method of assignment paternity based on probabilities, also called categorical attribution, where paternity is completely assigned to the supposed parent with the greatest probability (Flanagan & Jones, 2019). The likelihood is based on the observation of the genotypes, given the hypothesis, which is the proposed relationship between the individuals. In what probability is it calculated using the Mendelian inheritance rules (Marshall et al., 1998; Kalinowski et al., 2007). In the assignment of parents, the reason involves the hypothesis that the dyad (parent-offspring) or triad (sire-offspring-dam) represents a true set of parents and offspring versus the hypothesis that the individuals are not related.

In chapter two, the results demonstrate that the pedigree reconstructed by us had a high level of confidence (> 99.9%), and is similar to that used by the Brazilian Association of Zebu Breeders (ABCZ) with respect to recent relationships, while there was an increase in older relationships. These results demonstrate that the use of this methodology is efficient for breeds in wich there is no well-established pedigree, as well as breeds in where pedigree registrations have only recently started, or wild species.

Genomic data in population structure analysis

Numerous studies evaluating inbreeding, effective population size (N_e), genetic diversity and other important population parameters were generally done based on information from pedigree records (Malhado et al., 2010; Reis-Filho et al., 2010; Santana Jr et al., 2014). Currently, with the genomic era, it is possible to estimate these population parameters more accurately with information contained in the genotype of the animals, and numerous studies have been done with different breeds and species (Peripolli et al., 2018; Ospina et al., 2019; Ablondi et al., 2020).

In chapter three, we observed that there was a decrease in N_e based on LD, when calculated in the last 50 generations. The average generation interval found was equal to 9.08 years. The runs of homozygosity (ROH) analysis, an indicator of homozygous segments and, consequently, inbreeding, confirmed the results of N_e. The ROH analysis showed that the levels of inbreeding, especially those based on short ROH segments, were moderate to high, suggesting the presence of ancient bottlenecks in Gir cattle genome. Furthermore, ROH also indicated a lower percentage in longer classes, showing a small change in genetic variability during the last generations.

The results indicate that we need to continue monitoring conservation measures to minimize the loss of genetic diversity in the breed and to adopt breeding strategies that minimize inbreeding.

Practical application of population structure results in breeding programs

Farm animal genetic diversity is required to meet current production needs in various environments, allowing the perpetuation of animals resistant to diseases and extreme temperatures, and to facilitate rapid adaptation to changing breeding objectives (Notter, 1999). Some simple population parameters, such as inbreeding coefficient, heterozygosity, linkage disequilibrium, and effective population size, have an important impact on the genetic variability of a population (Reis-Filho et al., 2010). Regardless of whether the parameters are calculated based on the pedigree or the genotype, these parameters are used basically, to monitor genetic diversity and to elucidate important factors that affect the genetic evolution of the analyzed population (Valera et al., 2005).

According Notter (1999), the production efficiency in livestock is intricately linked to the use of several genetic compositions, but the intensive selection of sires is leading to greater genetic uniformity in intensively bred species. This uniformity is due to the increase in the frequency of favorable additive alleles and a consequent break in the homeostatic regulatory mechanisms established under natural selection in wild populations.

The ROH, inbreeding coefficients and effective population size of Gir cattle, found in chapter three, can be considered in the mating decisions in the breeding program. Adopting possible strategies to increase the N_e, and mating management that minimize inbreeding, and do not depend on the massive use of a few sires with high genetic values, maintaining genetic variability in future generations. Furthermore, looking for ways to reduce the long generation intervals, maximizing the genetic progress of the population.

General conclusion

This dissertation provides important information that can be added to the Brazilian Dairy Gir Breeding Program (PNMGL) to increase the accuracy and volume of records, in addition to assisting in the development of mating management strategies. The likelihood ratio approach, while offering redundant information with respect to recent pedigree information, information can be added from previously unknown relationships, increasing the depth of the pedigree. While the results of population structure analyzes based on genomic data become a guide to assist genetic improvement and increase the genetic diversity of the breed.

References

Ablondi, M., C. Dadousis, M. Vasini, S. Eriksson, S. Mikko, and A. Sabbioni. 2020. Genetic diversity and signatures of selection in a native italian horse breed based on SNP data. Animals, 10:1005. https://doi.org/10.3390/ani10061005.

Falconer, D. S., and T. F. C. Mackay. 1996. Small Populations: II Less simplified conditions. Pages 70–76 in Introduction to quantitative genetics, 4th edition. Longman Scientific and Technical, Essex, England.

Flanagan, S. P., and A. G. Jones. 2019. The future of parentage analysis: From microsatellites to SNPs and beyond. Mol. Ecol. 28:544–567. https://doi.org/10.1111/mec.14988.

Geldermann, H., U. Pieper, W. E. Weber. 1986. Effect of misidentification on the estimation of breeding value and heritability in cattle. Journal of Animal Science, v. 63, n. 6, p. 1759–1768. https://doi.org/10.2527/jas1986.6361759x.

Harlizius, B., M. S. Lopes, N. Duijvesteijn, L. H. P. van de Goor, W. A. van Haeringen, H. Panneman, S. E. F. Guimarães, J. W. M. Merks, E. F. Knol. 2011. A single nucleotide polymorphism set for paternal identification to reduce the costs of trait recording in commercial pig breeding, J. Anim. Sci. 89:1661–1668. https://doi.org/10.2527/jas.2010-3347.

Henderson, C. R. 1975. Best linear unbiased estimation and prediction under a selection model. Biometrics. 31:423–447. https://doi.org/10.2307/2529430.

Kalinowski, S. T., M. L. Taper, & T. C. Marshall. 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Mol. Ecol. 16:1099–1106. https://doi.org/10.1111/j.1365-294X.2007.03089.x.

Legarra, A., D. A. Lourenco, and Z. Vitezica. 2018. Bases for genomic prediction. Paris, França: INRA, Animal Genetics Department.

Lopes, P. S., R. de A. Torres, and A. V Pires. 2005. Selection methods. Pages 97–102 in Animal Breeding Theory. Belo Horizonte: FEPMVZ.

Malhado, C. H. M., P. L. S. Carneiro, A. C. M. Malhado, J. A. M. Martins, R. Martins Filho, and R. Bozzi. 2010. History of registered Gyr breed in Brazilian Northeast: population structure and genetic improvement of growth traits. Cienc. Rural, 40:1385–1391. https://doi.org/10.1590/S0103-84782010000600023.

Marshall, T. C., J. B. K. E. Slate, L. E. B. Kruuk, and J. M. Pemberton. 1998. Statistical confidence for likelihood-based paternity inference in natural populations. Mol. Ecol. 7:639–655. https://doi.org/10.1046/j.1365-294x.1998.00374.x.

Notter D. R. 1999. The importance of genetic diversity in livestock populations of the future, J. Anim. Sci. 77:61–69, https://doi.org/10.2527/1999.77161x.

Ospina A. M. T., A. M. Maiorano, R. A. Curi, G. L. Pereira, M. E. Zerlotti-Mercadante, J. N. dos S. Gonçalves Cyrillo, R. R. Aspilcueta-Borquis, and J. A. I. de V. Silva. 2019. Linkage disequilibrium and effective population size in Gir cattle selected for yearling weight. Reprod. in Domest. Anim., 54:1524–1531. https://doi.org/10.1111/rda.13559.

Patry, C., V. Ducrocq. 2011. Accounting for genomic pre-selection in national BLUP evaluations in dairy cattle. Genet. Sel. Evol. 43:30. https://doi.org/10.1186/1297-9686-43-30.

Peripolli, E., N. B. Stafuzza, D. P. Munari, A. L. F. Lima, R. Irgang, M. A. Machado, M. A., J. C. do C, Panetto, R. V. Ventura, F. Baldi, and M. V. G. B da Silva. 2018. Assessment of runs of homozygosity islands and estimates of genomic inbreeding in Gyr (*Bos indicus*) dairy cattle. BMC Genomics, 19:1–13. https://doi.org/10.1186/s12864-017-4365-3.

Reis-Filho, J. C., P. S. Lopes, R. D. S. Verneque, R. D. A. Torres, R. L. Teodoro, and P. L. S. Carneiro. 2010. Population structure of Brazilian Gyr dairy cattle. Rev. Bras. Zootec. 39:2640–2645. https://doi.org/10.1590/S1516-35982010001200012.

Ron, M., Y. Blanc, M. Band, E. Ezra, and J. I. Weller. 1996. Misidentification rate in the Israeli dairy cattle population and its implications for genetic improvement. J. Dairy Sci. 79:676-681. https://doi.org/10.3168/jds.S0022-0302(96)76413-5.

Santana Jr, M. L., R. J. Pereira, A. B. Bignardi, L. El Faro, H. Tonhati, and L. G. Albuquerque. 2014. History, structure, and genetic diversity of Brazilian Gir cattle. Livestock Science, 163:26–33. https://doi.org/10.1016/j.livsci.2014.02.007.

Senneke, S. L., M. D. MacNeil, and L. D. Van Vleck. 2004. Effects of sire misidentification on estimates of genetic parameters for birth and weaning weights in Hereford cattle. J. Anim. Sci., 82:2307–2312. https://doi.org/10.2527/2004.8282307x.

Valera, M., A. Molina, J. P. Gutiérrez, J. Gómez, and F. Goyache. 2005. Pedigree analysis in the Andalusian horse: population structure, genetic variability and influence of the Carthusian strain. Livest. Prod. Sci. 95:57–66. https://doi.org/10.1016/j.livprodsci.2004.12.004.

Van de Goor, L. H. P., M. T. Koskinen, and W. A. van Haeringen. 2011. Population studies of 16 bovine STR loci for forensic purposes. Int. J. Legal Med. 125:111–119. https://doi.org/10.1007/s00414-009-0353-8.

Van Vleck, L. D. 1970. Index selection for direct and maternal genetic components of economic traits. Biometrics. 26:477–483. https://doi.org/10.2307/2529103.