Genetic diversity analysis in dairy cows of Nariño, southwestern Colombia

Análise da diversidade genética em vacas leiteiras de Nariño, sudoeste da Colômbia

Maria Fernanda Betancur Zambrano[1]*; Juan Carlos Rincón Florez[2]; Rodrigo Ochoa[3]; Carlos Eugenio Solarte Portilla[4]

Highlights

Ancestry results confirmed Holstein contribution in crossbred animals in Nariño.
Genetic diversity within subpopulations showed average Ho of 0.40 and He of 0.36.
Fst results ranged from 0.01 to 0.47 with significant differences in some cases.

Abstract

In Colombia, different dairy breeds were introduced from Europe and the United States, which underwent different crossing and selection processes that generated specific qualities or differences and which likely have their own genomic structure. To characterize genetic diversity, population structure, and admixture, we used genotypes from 23,182 autosomal single nucleotide polymorphisms (SNPs) of 130 animals representing four dairy cattle breed groups from Nariño. In addition, we merged genotypes from 43,043 autosomal SNPs, from 137 animals from the Decker database (Decker et al., 2014) (DRYAD doi:10.5061/dryad.th092). After the quality control process of pruning the merged dataset, we were left with 7,475 autosomal SNPs shared by both databases of Nariño (127 samples) and Decker (135 samples). Genetic diversity levels were moderate in all breeds (average observed heterozygosity = 0.40). Based on the fixation index values, we conclude that Brahman individuals were more differentiated than the taurine breeds (-0.374 to 0.076 for Brown Swiss). Pairs between taurine breeds showed low genetic differentiation (0.011-0.479). Principal component analysis revealed that in both the Nariño and Decker databases, the taurine formed the most compact cluster compared with other breeds known not to share the same ancestry, and Jersey, Brown Swiss, and Normand individuals exhibited high similarity with Holstein individuals. Hierarchical cluster analysis with Admixture revealed that Brahman, Jersey, Normand, and Holstein from the Decker databases most of which were clustered together with the dairy breeds of the

[1] Ph.D. Student, Molecular Genetic Group, Biology Institute, Universidad de Antioquia, Medellín, Colombia 050010. E-mail: mafebeza@gmail.com

[2] Dr., Agricultural Sciences, Animal Sciences Department, Faculty of Agricultural Sciences, Universidad Nacional de Colombia, Sede Palmira. E-mail: jcrincon@unal.edu.co

[3] Dr., Chemical Sciences, Program for the Study and Control of Tropical Diseases, PECET, Universidad de Antioquia, Medellín, Colombia 050010. E-mail: rodrigo.ochoa85@gmail.com

[4] Dr., Veterinary Sciences, Animal Health and Production Group, Zootechnic Department, Faculty of Livestock Sciences, Universidad de Nariño. Pasto, Colombia 520002. E-mail: csolarte@hotmail.com

* Author for correspondence

Semina: Ciênc. Agrár. Londrina, v. 43, n. 6, p. 2563-2578, nov./dez. 2022

2563

Nariño highland tropics are not able to create different groups, thus having greater similarity with each other. This can be explained by the crosses made by farmers to increase milk production volume, always based on the Holstein breed with semen of bulls from America and Canada. Detrimental impacts due to intensive selection might cause some specific traits from the region to be fixed in the offspring, which can influence their adaptive capacity to the highland tropics*.

Key words: *Bos taurus.* Dairy yield. Polymorphism. SNP.

Resumo

Na Colômbia foram introduzidas diferentes raças leiteiras trazidas da Europa e dos Estados Unidos, que passaram por diferentes processos de cruzamento e seleção, gerando o desenvolvimento de características ou diferenças específicas nas raças bovinas e provavelmente, exibem sua própria estrutura genômica. Para caracterizar a diversidade genética e estrutura populacional foram utilizados genótipos de 23.182 polimorfismos de nucleotídeo único (SNPs) autossômicos de 130 animais, representando quatro grupos de raças de gado leiteiro de Nariño. Além disso, juntamos genótipos de 43.043 SNPs autossômicos, de 137 animais do banco de dados Decker (Decker et al., 2014) (DRYAD doi:10.5061/dryad.th092). Após o processo de controle de qualidade, o conjunto de dados misto ficou com 7.475 SNPs autossômicos compartilhados em ambos os bancos de dados de Nariño (127 amostras) e Decker (135 amostras). Os níveis de diversidade genética foram moderados em todas as raças (heterozigose média observada = 0,40). Com base nos valores do índice de fixação, concluímos que os indivíduos Brahman (zebuínos) foram mais diferenciados em comparação às raças taurinas (-0,374 a 0,076 para Pardo Suíço). Pares entre as raças taurinas apresentaram baixa diferenciação genética (0,011 – 0,479). O resultado dos componentes principais mostra que nos dois bancos de dados, Nariño e Decker, os taurinos formaram o cluster mais compacto em comparação com outras raças conhecidas por não compartilharem a mesma ancestralidade; indivíduos Jersey, Pardo Suíço e Normando denotaram alta similaridade com indivíduos Holandeses. A análise de agrupamento hierárquico com Admixture mostrou que Brahman, Jersey, Normando e Holandês da base de dados Decker, a maioria deles também agrupados com as raças leiteiras do alto trópico de Nariño, não são capazes de criar grupos diferentes, portanto, mostram maior semelhança entre eles. Isso pode ser explicado devido aos cruzamentos feitos pelos pecuaristas para aumentar o volume de produção de leite, sempre tendo como base a raça Holandesa com sêmen de touros da América e Canadá. Impactos prejudiciais, devido à seleção intensiva, podem fazer com que algumas características específicas da região sejam fixadas na prole e isso pode influenciar a sua capacidade adaptativa aos trópicos altos.

Palavras-chave: *Bos taurus.* Polimorfismo. Produção de leite. SNP.

Introduction

Holstein is the most common dairy breed in the world, and this is not different in Colombia, although other breeds are used on dairy farms (e.g. Jersey, Brown Swiss), besides dual-purpose breeds (e.g. Normand) (Ajmone-Marsan et al., 2010). Holstein was introduced into Colombia during the late 19th and early 20th centuries from Holland and North America (Rincon et al., 2018). Dairy farms are located in the highland tropics (>2000 m elevation, where weather is colder), which include the states of Antioquia, Cundinamarca,

2564

Semina: Ciênc. Agrár. Londrina, v. 43, n. 6, p. 2563-2578, nov./dez. 2022

and Nariño, where pastures of Kikuyu grass (Cenchrus clandestinum) predominate. With an average milk production of 12 to 14 kg/day in grazing with a stocking rate between one and two cows per hectare, Colombia produced 7,066 million liters of milk in 2017. This amount involves an estimated 99,000 producers in specialized dairy systems and 250,000 in dual-purpose systems (Carulla, 2016; Federacion Colombiana de Ganaderos [FEDEGAN], 2017). However, breeding programs in Colombia are new and sires are selected without appropriate selection criteria such as directed mating, genetic and reproductive evaluations, performance tests, genotyping analysis, or DNA sequencing, so it is difficult to predict the consequences (Rincon et al., 2018).

With the development of next-generation sequencing technologies (NGS), the annotation of single-nucleotide polymorphisms (SNPs) has been routinely employed to infer population history, genetic diversity, genome-wide association, and genome structure in domestic animals (Gautier et al., 2007; Decker et al., 2014; Mastrangelo et al., 2014; Edea et al., 2015; Šidlová et al., 2015). Among livestock species, this technology has been successfully applied in cattle thanks to many factors such as evolutionary history, genetic structure, and economic conditions, which render cattle particularly suitable for the application of genome-assisted selection (Nicolazzi et al., 2014). Genome-wide comparison of populations from different geographical regions can also detect potential candidate genes associated with ecological adaptations (Edea et al., 2015; Zambrano et al., 2020; Verardo et al., 2021). In livestock populations, approximately half of the genetic diversity is shared across breeds, while the other half is observed within single breeds (Sponenberg & Bixby, 2007).

Studies of genetic diversity are useful to understand the evolution of breeds, gene pool development, and the level of differentiation between breeds as well as to devise effective strategies for improving, managing, and conserving farm animal genetic resources (Melka & Schenkel, 2012; Edea et al., 2015). Numerous studies using various methods and types of molecular markers to characterize genetic diversity within and between breeds have been published (Decker et al., 2014; Lenstra et al., 2012). Genome-wide single nucleotide polymorphism (SNP) data have pushed the characterization of genetic diversity in cattle breeds forward (Signer-Hasler et al., 2017).

Despite the large phenotypic diversity between cattle populations in Colombia, particularly dairy cattle, there is a lack of information related to genetic diversity and relationships between them. Few studies on dairy using Restriction Enzyme Digestion in Polymerase Chain Reaction (PCR-RFLP) and single SNPs have been reported that describe the genetic diversity of the Holstein population of Antioquia and its structure (Rincon et al., 2011, 2013; Rodríguez et al., 2013; Isaza et al., 2017), or the genetic variability of Creole cattle through molecular techniques (Bejarano et al., 2012).

Dairy cattle in Nariño have been subjected to selection through breeding programs (Solarte-Portilla & Zambrano-Burbano, 2012; Zambrano et al., 2020). In addition, these cattle have undergone changes due mainly to indirect selection by the use of reproductive biotechnology

(artificial insemination and embryo transfer), generating technical and economic dependence on other countries (e.g. United States, Canada, and some European countries). The interaction of these factors with the environment of the region may have influenced the adaptation of the animals to the mountain conditions typical of the highland tropics of Nariño. Therefore, additional studies to determine the effect generated by such factors are required in the highland tropics. Here we assessed the status of genetic diversity, population structure, and admixture within and between three pure cattle populations from Nariño using information from 23,182 autosomal SNPs.

Material and Methods

Animals and genotyping

The animals under study belong to a project called Selection Research Using Genomic and Polygenic Models for the Genetic Improvement of Dairy Cattle in the Highland Tropics of Nariño, financed with resources of Sistema General de Regalías (SGR) in cooperation with Cooperativa de Lacteos de Nariño and executed by the Animal Production Research Group of Universidad de Nariño, cod BPIN No. 2013000100091. Genotypic data were collected from 37 herds in the first, second, and third districts. Table 1 describes the milk districts of Pasto, Pupiales, and Guachucal.

A total of 130 individuals born in Nariño, southwestern Colombia, were studied. These animals were chosen randomly and proportionally to herds that were enrolled in the Nariño breeding program. They represented four cattle breed groups among which are Holstein (HOL), Brown Swiss (BSW), Jersey (JER), and Normand (NOR), from unrelated dams and sires (Table 2). Blood samples were collected following the ethical guidelines established by the Animal Welfare Committee at the University of Nariño. DNA was extracted from blood using an Ultraclean BloodSpin kit (MoBio Laboratories Inc, Carlsbad, California, USA). DNA was stored at -20 °C until genotyping. Total DNA from selected individuals was genotyped using the GeneSeek Genomic Profiler LD v3 (23,182 SNPs) (http://www.neogen.com/GeneSeek/). In addition, samples of animals analyzed in previous studies worldwide were considered for this study. Using genotypes from 43,043 autosomal single nucleotide polymorphism markers, we merged data of 137 animals from the Decker database (Decker et al., 2014) (DRYAD doi:10.5061/dryad.th092, see Table 2): 28 JER, 30 NOR, 30 HOL, 24 BSW, and 25 Brahman (BRM). The Brahman were chosen because they would represent a differentiation between Bos taurus and Bos indicus. To make a dissimilarity in the nomenclatures of the breeds from the two databases, we will add a w at the end of the worldwide samples in the results of the analyses.

2566

Semina: Ciênc. Agrár. Londrina, v. 43, n. 6, p. 2563-2578, nov./dez. 2022

Table 1
Geographical coordinates and main climatic characteristics of the three milk districts of the department of Nariño, Colombia

District	Coordinates	Municipalities	Height msnm	Rainfall Mm/yr.	Pluv (mm/yr.)	Temp ° c	Light (hours)
Pasto	1°, 12`, 41`` North latitude 77°, 16`, 52`` West longitude	Pasto, Tangua, Yacuanquer, Nariño and Buesaco	2,527	960	771.6	13	1,142.5
Pupiales	0°, 52`, 21`` North latitude 77°, 38`, 34`` West longitude	Pupiales, Ipiales, Aldana, Carlosama and Gualmatán,	2,900	960	853.3	11.03	1,516.8
Guachucal	0°, 57`, 50`` North latitude 77°, 44`, 04`` West longitude	Guachucal, Cumbal, Sapuyes and Túquerres	3,087	940	959.0	10.8	1,309.2

Source: Research group animal production and health. Line of Genetics and Animal Improvement. Universidad de Nariño. Pasto-Colombia (2009).

Table 2
Provenance for all samples included in the analyses and results of statistical population analysis

Abb	Name	Tax	D	N	Ho	s.d	He	s.d	Fis	p-value
HOLw	Holstein	*Bos t. taurus*	1	30	0.428	0.130	0.427	0.102	0.047	0.394
JERw	Jersey	*Bos t. taurus*	1	28	0.383	0.155	0.380	0.132	-0.072	0.578
NORw	Normande	*Bos t. taurus*	1	30	0.426	0.129	0.412	0.111	0.013	0.773
BWSw	Brown Swiss	*Bos t. taurus*	1	24	0.411	0.154	0.396	0.124	-0.026	0.456
BRMw	Brahman	*Bos t. indicus*	1	25	0.288	0.166	0.296	0.157	-0.374	0.340
HOL	Holstein	*Bos t. taurus*	2	60	0.433	0.092	0.432	0.089	0.059	0.621
JER	Jersey	*Bos t. taurus*	2	18	0.407	0.159	0.402	0.120	-0.011	0.502
NOR	Normande	*Bos t. taurus*	2	46	0.420	0.126	0.415	0.108	0.020	0.626
BWS	Brown Swiss	*Bos t. taurus*	2	6	0.464	0.204	0.440	0.112	0.076	0.732

Samples of the two databases (Decker and Nariño). Abb: letter code used to identify the breed throughout the manuscript; Name: breed name; Tax: taxonomic status indicating the species name; D: data set's origin, Decker et al. (2014), (1), Nariño (2); N: breed's sample size; Ho: observed heterozygosity; He: expected heterozygosity; s.d: standard deviation; Fis: inbreeding coefficient.

Semina: Ciênc. Agrár. Londrina, v. 43, n. 6, p. 2563-2578, nov./dez. 2022

2567

Quality control of SNP data

Data quality control for the Nariño database was performed using PLINK v1.9 software (Purcell et al., 2007). Only SNPs located on autosomes were considered for further analyses, and SNPs assigned to the X chromosome were excluded. Monomorphic SNPs with more than 10% of missing genotypes were discarded from the analysis. All SNPs coordinates were converted to UMD3.1 bovine assembly (RefSeq: GS-3155939409), SNPchiMp database (Nicolazzi et al., 2014). SNPs from relevant dairy breeds were extracted and analyzed together. Individuals missing more than 5% of genotypic data were discarded (sample call rate < 95%). SNPs with a call rate below 95% were removed from further analysis. Autosomal SNPs shared (7,475) in both databases of Nariño (127 samples) and Decker (135 samples) were used for genetic relationship and ancestry proportion estimations. Afterwards, the quality control process of pruning was performed on data using the –indep-pairwise (50 10 0.2) option. Pruning was proceeded by calculating linkage disequilibrium (LD) for 50 marker sliding windows, with a new window obtained by shifting 10 markers along the length of the chromosome. Marker pruning was carried out when the LD between a pair of markers was 0.2 (Calus & Vandenplas, 2018; Chagunda et al., 2018). This resulted in 262 samples, of which 6,080 SNPs were used for genetic diversity analysis.

Genetic diversity

Along with R software (https://www.r-project.org/), PLINK v1.90 (Purcell et al., 2007) was used to generate population parameters. The dataset was processed for genetic diversity parameters that included expected heterozygosity (He) and observed heterozygosity (Ho) levels as well as inbreeding (F) for each population. Pairwise distances (Fst) (Weir & Cockerham, 1990) and Slatkin's distance measure were analyzed for the estimation of genetic distance and genomic relationship between samples using Arlequin 3.5.2.2 (Excoffier et al., 2005).

Genetic population relationships

Principal component analysis (PCA) was prepared for each dataset for stratification of different populations along different plot coordinates. The binary files (.bed, .bim, and .fam files) were generated for datasets using the PLINK program (Purcell et al., 2007). Eigenvectors were calculated with the aid of packages in R software (https://www.r-project.org/) and then plotted on the PCA plot.

The pruned dataset was processed in the ADMIXTURE program (Alexander et al., 2009) for ancestry and admixture analysis using the model-based approach from the autosomal SNP panel of dairy cattle populations of Nariño with other dairy cattle populations, to identify the main contribution of the breed to its formation. The maximum likelihood estimation (MLE) approach was used and datasets were processed for different K values (K value: parameter describing the number of subpopulations that make up the total population in a dataset provided as input). The K value (as input to the program) ranged from 1 to 20 per dataset in the ADMIXTURE program. The K value with minimum cross-validation error was assessed

2568

Semina: Ciênc. Agrár. Londrina, v. 43, n. 6, p. 2563-2578, nov./dez. 2022

independently for both datasets. Cross-validation error values helped to decide the best K value of each dataset.

Results and Discussion

The following results considered the combination of the two databases: Nariño (130 samples) and Decker (137 samples), which shared 7,475 autosomal SNPs.

Genetic diversity within subpopulations was accounted for as a measure of heterozygosity (observed and expected). The genetic diversity of a species allows them to respond and adapt, or not, to traits or changes in the environment. This is done at the chromosomal level, where recombination or mutation is carried out gradually and can give better or worse adaptive traits to the following generations (Black IV et al., 2001). After quality control and pruning on each population, the remaining data for diversity analysis were 6,808 SNPs from 262 animals representing several breeds. The average Ho value found for the whole sample was 0.407, and the average He was 0.36019 (Table 2). Taurine breeds showed greater diversity than the zebu (HOL: 0.433-0.432; HOLw: 0.428-0.427; JER: 0.407-0.402; JERw: 0.383-0.380; NOR: 0.420-0.415; NORw: 0.426-0.412; BSW: 0.464-0.441; BSWw: 0.411-0.396; and BRMw: 0.288-0.296). In the BRMw breed, He was greater than observed, but the values were not representative.

In Nariño, no diversity analysis has been done with SNPs approaches; the only reference of contrast is the analysis carried out by Mejía et al. (2015), who used 11 microsatellite loci in the same breeds but different samples and found average Ho

values greater than 0.6 for all breeds. These reflect a significantly high amount of genetic diversity in the population. In the same way, little information is currently available to compare different cattle populations from Colombia. Authors have reported results using other approaches, primarily in Holstein (Rincon et al., 2013; Rodríguez-Ramilo et al., 2015; Madrid et al., 2015; Isaza et al., 2017; Rincon et al., 2018). As a counterpart, studies based on the Illumina SNP chip reported diversity expressed in terms of Ho and He over all loci in Jersey and Holstein Friesian bulls: 0.26-0.36 and 0.31-0.42, respectively (Melka & Schenkel, 2012); and in Holstein from Australia: 0.33-0.34 (Zenger et al., 2007). The first values represent a Fis higher than that found in our study (Fis value of Gyr and Holstein from Decker database: 0.020 and 0.010; Jersey and Holstein from Nariño: -0,027 and -0,013, respectively), which means that our population has less diversity than Melka's, suggesting a possible adaptation to the local conditions of the Colombian tropics. We could also argue that having a population size reduced in diversity can increase reproduction between related individuals, which leads to high inbreeding coefficients in the evaluated population (Black IV et al., 2001).

In livestock species, heterozygote deficiencies can be explained by several factors, such as selection, population subdivision, drift, and inbreeding (Melka & Schenkel, 2012). It could be because populations have a degree of inbreeding (pairings within the group) and due to an intensive selection process to enhance production traits and improve herd economy (Suhada et al., 2016). The approaches described above allowed the identification

Semina: Ciênc. Agrár. Londrina, v. 43, n. 6, p. 2563-2578, nov./dez. 2022

2569

of distinct breeds, yet without determining their differences. For practical utility, diversity between breeds is favorable, as it allows optimizing targeted mating programs and creating hybrid populations that offer tools necessary for good decision making in the selection process for the dairy population of the Nariño tropics.

Fixation indexes (Fis, Fit, and Fst) are the most widely used parameters for studying the genetic differentiation of populations. These indices have been originally defined in terms of the correlations of two uniting gametes (Wright, 1965). Table 2 shows the Fis values found in our study. The samples were in a range of values from -0.374 for BRMw to 0.076 for BSW. Pairwise Slatkin's Fst (Excoffier et al., 2005) was used to study population differentiation. A higher Fst value indicates the presence of genetic differentiation between subpopulations, which implies that pairs of genes between individuals within subpopulations are more related than those of individuals between subpopulations (Melka & Schenkel, 2012).

Pairwise Slatkin's Fst values ranged between 0.011 and 0.479, with statistically significant differences in some cases (Table 3). Pairwise Fst estimates revealed the lowest level of genetic differentiation between Holstein cattle from Decker and Holstein from Nariño (Fst = 0.011), denoting their close genetic relationship; and the greatest divergence between Brahman (indicus) and Jersey from Decker (taurus) (Fst = 0.479). As expected, Brahman individuals were more differentiated in comparison to the taurine breeds (0.479-0.290). Pairwise estimates between taurine breeds in different locations were lower: JERw – JER (0.023), NORw – NOR (0.018), and BSWw – BSW (0.063). Genetic differentiation between Decker's taurine breeds ranged from 0.157 between BSWw – HOLw, 0.181 between JERw – HOLw, to 0.180 between JERw – NORw; and was even smaller when only samples of breeds from the Nariño region were compared: JER – NOR (0.115), JER – HOL (0.093), JER – BSW (0.071), and HOL – NOR (0.084). Other studies in which comparisons were made between Holstein and other dairy breeds found Fst values of 0.14 between HOL – JER and 0.15 between HOL – BWS (taurine Brown Swiss), which are higher than the pairwise pairwise of Nariño samples and similar to those of Decker samples (Melka & Schenkel, 2012; Signer-Hasler et al., 2017).

2570

Semina: Ciênc. Agrár. Londrina, v. 43, n. 6, p. 2563-2578, nov./dez. 2022

Table 3
Slatkin's genetic distance between subgroups or populations

Breeds[1]	BRMw	BSW	BSWw	HOL	HOLw	JER	JERw	NOR	NORw
BRMw	0.000								
BSW	0.439	*0.000*							
BSWw	0.451	*0.063*	0.000						
HOL	0.290	*0.053*	0.134	0.000					
HOLw	0.369	*0.065*	0.157	0.011	0.000				
JER	0.427	*0.071*	0.169	0.093	0.111	0.000			
JERw	0.479	*0.143*	0.223	0.149	0.181	0.023	0.000		
NOR	0.359	*0.091*	0.146	0.084	0.100	0.115	0.164	0.000	
NORw	0.384	*0.100*	0.154	0.094	0.109	0.128	0.180	0.018	0.000

[1] Brahman Decker (BRMw), Brown Swiss Nariño (BSW), Brown Swiss Decker (BSWw), Holstein Nariño (HOL), Holstein Decker (HOLw), Jersey Nariño (JER), Jersey Decker (JERw), Normand Nariño (NOR) and Normand Decker (NORw).

The Holstein breed is a cosmopolitan population and, as suggested by Doormaal et al. (2005), a global effort would be needed to achieve this goal, given that in some countries groups of animals may exist that are less related to those used globally. That would be the reason why the group of individuals from Decker's database and the individuals of European studies may have not been affected by crossing processes with other breeds and considered pure, while cattle from Nariño have a different management. Farmers in this Colombian region are not used to keeping genealogical, productive, or reproductive individual records. Moreover, they make crosses aiming to increase production without any knowledge about improvement, and in most cases, this process is done empirically. This could suggest the existence of inbreeding between Nariño breeds, since genetic material (imported semen) is used for directed crosses in most production systems from the region, which is mainly based on Holstein bulls to increase the volume of milk per day. In addition, this reveals no genetic differentiation between the subpopulations of Nariño samples that can be considered as part of a large population, where random mating occurs.

After quality control and pruning on each population, the data subjected to PCA were 6,808 SNPs from 262 animals representing several breeds. Results of PCA of both Nariño's and Decker's SNP datasets clearly grouped the populations by ancestry (Figure 1). The combination of the first two principal components (PC), namely, PC1 and PC2, grouped the samples according to their ancestral origin. PC1 accounted for 23.02% of the total variation and separated the B. indicus (BRM) cattle cluster from another B. taurus cattle cluster (HOL, JER, NOR, and BSW); PC2 explained 16.36% of the total variation and separated the animals in the cluster by their breed composition and between the samples of the Colombian and Decker B. taurus clusters. The Nariño group split the data on the first axis according to

Semina: Ciênc. Agrár. Londrina, v. 43, n. 6, p. 2563-2578, nov./dez. 2022

2571

the genetic distance of taurine breeds, which formed clusters almost overlapping each other due to genetic relatedness. Results of PCA of both Nariño's and Decker's SNP datasets allowed us to identify two main groups, separated by the geographical origin of the taurus and indicus subspecies. The Holstein were slightly separated from the rest of the taurine breeds; Normand, Jersey, and Brown Swiss formed a compact cluster compared to Nariño taurine breeds known for sharing the same ancestry; some NOR, JER, and BSW from Decker's database displayed high similarity with NOR, JER, and BSW individuals from Nariño. The results also revealed differences between the taurine and zebu breeds, which, as expected, show ancestral and demographic differentiation. These results suggest that thanks to the selection to which the European breeds have been subjected in this region, there may be a differentiation stemming from random mating, giving the hardy Creole animals greater adaptation to the region, which leads to more efficient production and favors the economy of producers.
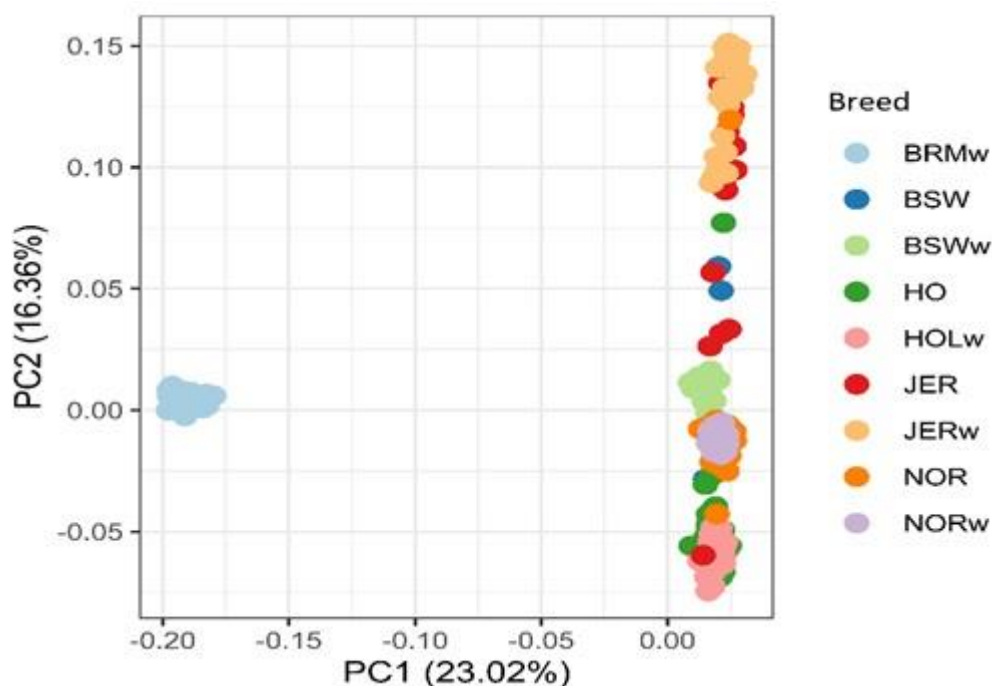

Figure 1. Principal components analysis (PCA) of 262 individuals based on 7,080 SNPs shared in dairy cattle populations from Nariño – Colombia and Decker's databases. PC1 accounts for 23,02% of the total variation; PC2 accounts for 16,36% of the total variation. B. indicus (Brahman BRM) B. taurus (Holstein HOL, Brown Swiss BSW, Jersey JER, and Normand NOR), individuals with the w indicate that they are from the Decker's database.

2572

Semina: Ciênc. Agrár. Londrina, v. 43, n. 6, p. 2563-2578, nov./dez. 2022

Production systems in southwestern Colombia are mostly smallholdings, with each farm having a maximum of 50 animals, and no scientific strategy is adopted to improve production. On these farms, few births are properly registered, since most farmers are not used to keeping reproductive or sanitary records. When cows reach sexual maturity, crosses are expected to be made to improve a productive trait. In most cases, crosses are made with international semen, but at times with national bulls that can be unregistered or that phenotypically appear like Holstein but are not indeed registered as such. In this way, we observe that there is a high possibility of poor handling of records, as shown in Figure 1, where we find animals registered as Jersey, Brown Swiss, or Normand that are actually in the Holstein breed cluster. Considering that keeping records on the cattle is not yet a common practice in Colombia, one must be careful when making assertions about their origin. Traditional production systems lacking individual records of origin result in a mixture of breeds commonly named "-like". Thus, the phenotype is closer to Holstein for the majority of individuals in the population and the SNP frequencies confirm this observation.

Hierarchal clustering of the dataset was performed as recommended in Admixture software, running for larger values of K (e.g., 20) (Figure 2). Data were plotted in order of the fraction of ancestry they shared with other breeds. We found that K = 5 was the optimum number of clusters in the sample, at which breeds started to show differentiation (Figure 3). The genetic ancestry observed in the Nariño dairy cattle population corresponds to that observed in Holstein individuals by Decker et al. (2014). Results showed that major taurine and zebu groups showed large

disparity; BRM, NORw, JERw, and HOLw breeds are able to create different groups, most of them overlapping each other, even though some subgroup structure within all breeds may explain this trend, similarly to Edea et al. (2015). For example, some animals from the Nariño group at K = 4 can be recognized as Holstein individuals while they are in the Normand cluster, and animals that apparently fit more to the color palette of Jersey but are in the Holstein breed cluster. This continues to show the poor quality of records as well as the amount of empirical mating that farmers have made. Nonetheless, we must also recognize that there is a difference between HOL, JER, and NOR individuals from Decker and Nariño, which could be due to the random mating of Nariño cows with semen from American or Canadian bulls, which in turn have been the product of crossing aimed at improving their production and adaptability to tropical climates, possibly improving traits of economic interest. Therefore, there may be a change in the genetic structure of the Nariño animals that favors their adaptation to the mountain tropics to improve production.

These results, coupled with the previous ones about differentiation, show that the dairy cattle population in Nariño consists is mainly of Holstein individuals, reaching almost 85% of the population included in the genetic improvement program in southwest Colombia over time (Solarte-Portilla et al., 2009). The Holstein, introduced over four decades ago into Nariño, has allowed it to become the most important nucleus of the breed for the region, and some crosses with this breed have displayed economic merit comparable to that of pure Holstein. In the same way, ancestry results confirmed the important contribution of Holstein cattle to crossbred animals, as observed in PCA.
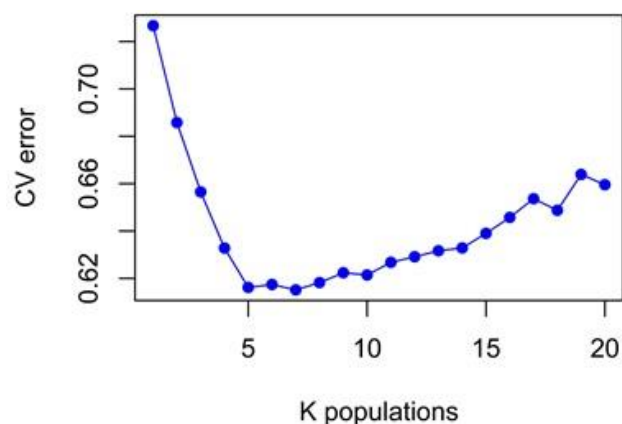
Semina: Ciênc. Agrár. Londrina, v. 43, n. 6, p. 2563-2578, nov./dez. 2022

2573

Figure 2. Cross-validation (CV) statistic values used to choose the most suitable number of clusters, K=20.
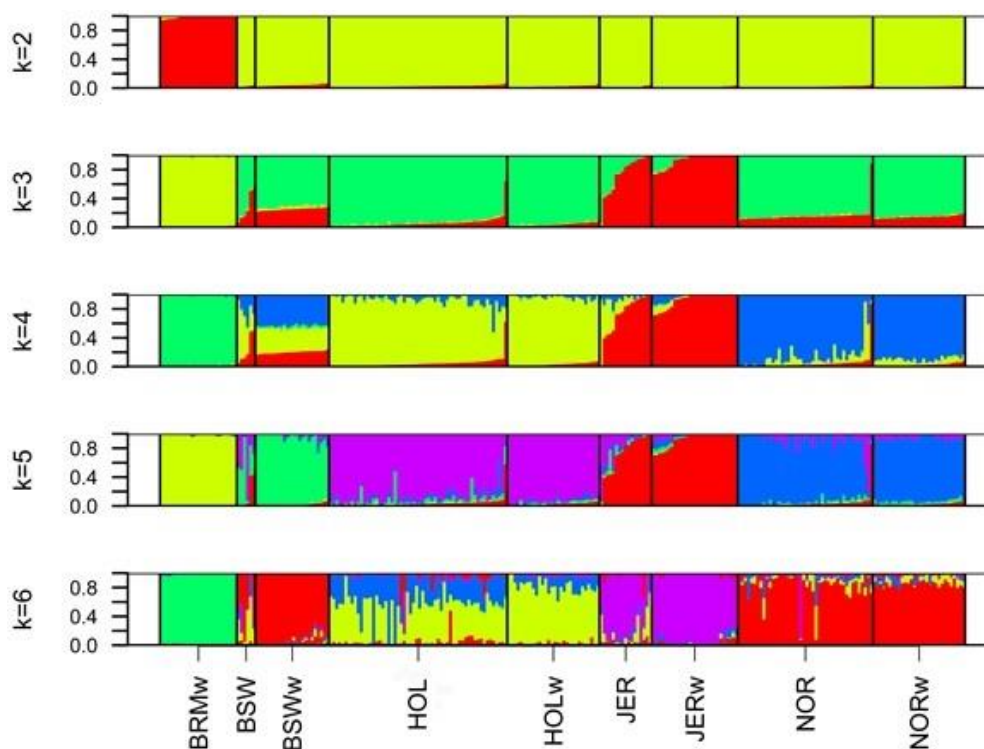

Figure 3. Analysis of Population Structure assignment for 262 individuals based on 7,080 SNPs shared in dairy cattle populations from Nariño – Colombia and Decker's databases using STRUCTURE (K= 2 to 6) and plotted with R software package. The amount of a color reflects the individual's proportion of genetic variation originating in the cluster of that color. Each breed is separated by a straight bar that is colored and labeled on the bottom of the plot. *Bos indicus* represents worldwide breeds with zebu ancestry, Brahman (BRMw), *Bos taurus* represent worldwide breeds with European taurine ancestry, Brown Swiss (BSWw), Holstein (HOLw), Jersey (JERw), Normand (NORw), and Nariño individuals Brown Swiss (BSW), Holstein (HOL), Jersey (JER) and Normand (NOR). For K=2, red descent represents *Bos indicus* ancestry, and yellow descent represents Bos taurus ancestry.

Conclusion

Our study is the first comprehensive analysis of dairy cattle with genome-wide genetic markers to understand the genetic diversity of cattle in the highland tropics of Nariño. Although we used only a few representative dairy breeds to understand the genetic backgrounds of these cattle, this study contributes to the knowledge of genetic diversity of local breeds in this region. After a long period during which mating processes were carried out with both imported and national semen, it is possible that some specific traits of the region were fixed in the offspring, causing differences between Holstein from the highland tropics of Nariño and the rest of the world. The implementation of genomic evaluation programs requires, at first, knowledge of the evolutionary history of the population, making it possible to optimize programs and understand the genetic resource available. For this reason, it is important that further research be carried out exploring the population structure, detecting markers associated with quantitative trait loci (QTLs) of adaptation to highland tropical conditions, as well as identifying regions of the genome with selection signals that may reveal mountain traits in animals or their association with traits of economic interest such as reproduction, production, and adaptability.

Funding acknowledgement

Acknowledgments

References

Ajmone-Marsan, P., Garcia, J. F., & Lenstra, J. A. (2010). On the origin of cattle: how aurochs became cattle and colonized the world. *Evolutionary Anthropology: Issues, News, and Reviews, 19*(4), 148-157. doi: 10.1002/evan.20267

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research, 19*(9), 1655-1664. doi:10.1101/gr.094052.109

Bejarano, D., Pedraza, A., Rocha, F., & Martinez, R. (2012). Variabilidad genética en subpoblaciones comerciales de la raza criolla colombiana Romosinuano. *Genética, Reproducción y Mejoramiento Animal, 13*(1), 97-107. doi: 10.21930/rcta.vol13_num1_art:24

Black IV, W. C., Baer, C. F., Antolin, M. F., & DuTeau, N. M. (2001). Population genomics: genome-wide sampling of insect populations. *Annual Review of Entomology, 46*(1), 441-469. doi: 10.1146/annurev.ento.46.1.441

Calus, M. P. L., & Vandenplas, J. (2018). SNPrune: an efficient algorithm to prune

large SNP array and sequence datasets based on high linkage disequilibrium. *Genetics Selection Evolution, 50*(34), 1-11. doi: 10.1186/s12711-018-0404-z

Carulla, J. E. (2016). Sistemas de producción lechera en Colombia : retos y oportunidades dairy production systems of Colombia: challenges and opportunities. *Archivos Latinoamericanos de Producción Animal, 24*(2), 9-13. https://ojs.alpa.uy/index.php/ojs_files/article/view/2526

Chagunda, M. G. G., Mujibi, F. D. N., Dusingizimana, T., Kamana, O., Cheruiyot, E., & Mwai, O. A. (2018). Use of high density Single Nucleotide Polymorphism (SNP) arrays to assess genetic diversity and population structure of dairy cattle in smallholder dairy systems: the case of girinka programme in rwanda. *Frontiers in Genetics, 9*(438), 1-8. doi: 10.3389/fgene.2018.00438

Decker, J. E., McKay, S. D., Rolf, M. M., Kim, J. W., Molina Alcalá, A., Sonstegard, T. S., Hanotte, O., Götherström, A., Seabury, C. M., Praharani, L., Babar, M. E., Correia de Almeida Regitano, L., Yildiz, M. A., Heaton, M. P., Liu, W. S., Lei, C. Z., Reecy, J. M., Saif-Ur-Rehman, M., Schnabel, R. D., & Taylor, J. F. (2014). Worldwide patterns of ancestry, divergence, and admixture in domesticated cattle. *PLoS Genetics, 10*(3), doi: 10.1371/journal.pgen.1004254

Doormaal, B. J. Van, Miglior, F., Kistemaker, G., & Brand, P. (2005). *Genetic diversification of the holstein breed in Canada and internationally.*

Edea, Z., Bhuiyan, M. S. A., Dessie, T., Rothschild, M. F., Dadi, H., & Kim, K. S. (2015). Genome-wide genetic diversity, population structure and admixture analysis in African and Asian cattle breeds. Animal: *An International Journal of Animal Bioscience, 9*(2), 218-226. doi: 10.1017/S1751731114002560

Excoffier, L., Laval, G., & Schneider, S. (2005). Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionry Bioinformatics, 1*(1), 47-50. doi: 10.1111/j.1755-0998.2010.02847.x

Federacion Colombiana de Ganaderos (2017). *Balance y perspectivas del sector Ganadero Colombiano 2016-2017.* FEDEGAN. https://doi.org/www.fedegan.org.co

Gautier, M., Faraut, T., Moazami-Goudarzi, K., Navratil, V., Foglio, M., Grohs, C., Boland, A., Garnier, J. G., Boichard, D., Lathrop, G. M., Gut, I. G., & Eggen, A. (2007). Genetic and haplotypic structure in 14 European and African cattle breeds. *Genetics, 177*(2), 1059-1070. doi: 10.1534/genetics.107.075804

Isaza, J. V., Lopez, A., & Echeverri, J. (2017). Análisis poblacional del ganado Holstein de Antioquia a partir de un polimorfismo del gen POU1F1. *Archivos de Zootecnia, 66*(254), 287-299. doi: 10.21071/az.v66i254.2324

Lenstra, J. A., Groeneveld, L. F., Eding, H., Kantanen, J., Williams, J. L., Taberlet, P., Nicolazzi, E. L., Sölkner, J., Simianer, H., Ciani, E., Garcia, J. F., Bruford, M. W., Ajmone-Marsan, P., & Weigend, S. (2012). Molecular tools and analytical approaches for the characterization of farm animal genetic diversity. *Animal Genetics, 43*(5), 483-502. doi: 10.1111/j.1365-2052.2011.02309.x

2576

Semina: Ciênc. Agrár. Londrina, v. 43, n. 6, p. 2563-2578, nov./dez. 2022

Madrid, S., Lopez, A., & Echeverri, J. (2015). INHA A192G polymorphism and its association with dairy traits in Antioquia Holstein cattle. *Archivos de Zootecnia, 64*(246), 147-154. https://www.redalyc.org/articulo.oa?id=49545650008

Mastrangelo, S., Saura, M., Tolone, M., Salces-Ortiz, J., Di Gerlando, R., Bertolini, F., Fontanesi, L., Sardina, M. T., Serrano, M., & Portolano, B. (2014). The genome-wide structure of two economically important indigenous Sicilian cattle breeds. *Journal of Animal Science, 92*(11), 4833-4842. doi: 10.2527/jas.2014-7898

Mejía, L. G., Hernández, R. A., Rosero, C. Y., & Solarte-Portilla, C. E. (2015). Análisis de la diversidad genética de ganado bovino lechero del trópico alto de Nariño mediante marcadores moleculares heterólogos de tipo microsatélite. *Revista de La Facultad de Medicina Veterinaria y de Zootecnia, 62*(3), 18-33. doi: 10.15446/rfmvz.v62n3.54938

Melka, M. G., & Schenkel, F. S. (2012). Analysis of genetic diversity in Brown Swiss, Jersey and Holstein populations using genome-wide single nucleotide polymorphism markers. *BMC Research Notes, 5*(161), 1-8. doi: 10.1186/1756-0500-5-161

Nicolazzi, E. L., Picciolini, M., Strozzi, F., Schnabel, R. D., Lawley, C., Pirani, A., Brew, F., & Stella, A. (2014). SNPchiMp: a database to disentangle the SNPchip jungle in bovine livestock. *BMC Genomics, 15*(123), 1-6. doi: 10.1186/1471-2164-15-123

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics, 81*(3), 559-575. doi: 10.1086/519795

Rincon, J. C., Lopez, A. H., & Echeverri, J. Z. (2011). Variabilidad en dos SNP´s localizados en dos secuencias codificantes. *Actas Iberoamericanas de Conservación Animal, 1*(1), 154-156. https://www.uco.es/conbiand/aica/templatemo_110_lin_photo/articulos/2011/Rincon2011_1_154_156.pdf

Rincon, J. C., Lopez, A. H., & Echeverri, J. Z. (2013). Estructura y diversidad genética en vacas holstein de antioquia usando un polimorfismo del gen bGH. *Revista MVZ Cordoba, 18*(1), 3346-3354. doi: 10.21897/rmvz.197

Rincon, J. C., Lopez, A., & Echeverri, J. (2018). Identifying signatures of recent selection in Holstein cattle in the tropic. *Revista Colombiana de Ciencias Pecuarias, 31*(1), 45-58. doi: 10.17533/udea.rccp.v31n1a06

Rodríguez, N. C., López, A. H., & Echeverri, J. Z. (2013). Estructura genética poblacional del gen lactoferrina bovino en vacas holstein del departamento de antioquia. *Revista MVZ Cordoba, 18*(1), 3355-3361. doi: 10.21897/rmvz.198

Rodríguez-Ramilo, S. T., Fernández, J., Toro, M. A., Hernández, D., & Villanueva, B. (2015). Genome-wide estimates of coancestry, inbreeding and effective population size in the spanish holstein population. *PLoS ONE, 10*(4), e0124157. doi: 10.1371/journal.pone.0124157

Semina: Ciênc. Agrár. Londrina, v. 43, n. 6, p. 2563-2578, nov./dez. 2022

2577

Šidlová, V., Moravčíková, N., Trakovická, A., Ferenčaković, M., Curik, I., & Kasarda, R. (2015). Production type of Slovak Pinzgau cattle in respect of related breeds. *Acta Fytotechnica et Zootechnica, 2015*(2), 25-29. doi: 10.15414/afz.2015.18.02.25

Signer-Hasler, H., Burren, A., Neuditschko, M., Frischknecht, M., Garrick, D., Stricker, C., Gredler, B., Bapst, B., & Flury, C. (2017). Population structure and genomic inbreeding in nine Swiss dairy cattle populations. *Genetics Selection Evolution, 49*(83), 1-13. doi: 10.1186/s12711-017-0358-6

Solarte-Portilla, C. E., & Zambrano-Burbano, G. L. (2012). Characterization and genetic evaluation of Holstein cattle in Nariño, Colombia. *Revista Colombiana de Ciencias Pecuarias, 25*(4), 539-547. doi: http://www.redalyc.org/articulo.oa?id=295024922002

Solarte-Portilla, C. E., Rosero-Galindo, C. Y., Burgos-Paz, W. O., & Caviedes-Castro, I. F. (2009). *Caracterización y Evaluación Genética de la Población Bovina Lechera del Trópico Alto de Nariño para la Conformación de Núcleos de Selección.* doi: 10.13140/2.1.3396.1762

Sponenberg, D. P., & Bixby, D. E. (2007). *Managing breeds for a secure future: strategies for breeders and breed associations.*

Suhada, H., Anwar, S., Maulana, H., & Yurnalis, D. (2016). Diversity of growth hormone gene and its relation with average daily gain in Simmental cattle in West Sumatera Province, Indonesia. *African Journal of Biotechnology, 15*(45), 2565-2571. doi: 10.5897/AJB2016.15553

Verardo, L. L., Silva, F. F. e, Machado, M. A., Carmo Panetto, J. C. do, Lima Reis Faza, D. R. de, Otto, P. I., Almeida Regitano, L. C. de, Silva, L. O. C. da, Egito, A. A. do, Socorro Maués Albuquerque, M. do, Zanella, R., & Silva, M. V. G. B. da. (2021). Genome-wide analyses reveal the genetic architecture and candidate genes of indicine, taurine, synthetic crossbreds, and locally adapted cattle in Brazil. *Frontiers in Genetics, 12*, e702822.1-15. doi: 10.3389/fgene.2021.702822

Weir, B. S., & Cockerham, C. (1990). *Genetic data analysis II: methods for discrete population genetic data.* Sinuer Assoc.

Wright, S. (1965). The Interpretation of population structure by f-statistics with special regard to systems of mating. *Evolution, 19*(3), 395-420. doi: 10.2307/2406450

Zambrano, M. F. B., Flórez, J. C. R., Ríos, A. C. H., Portilla, C. E. S., & Jesús Bedoya Berrio, G. de. (2020). Evaluation of runs of homozygosity and genomic inbreeding in Holstein cattle from Colombia. *Semina: Ciências Agrárias, 41*(6), 3397-3418. doi: 10.5433/1679-0359.2020v41n6Supl2p3397

Zenger, K. R., Khatkar, M. S., Cavanagh, J. A. L., Hawken, R. J., & Raadsma, H. W. (2007). Genome-wide genetic diversity of Holstein Friesian cattle reveals new insights into Australian and global population variability, including impact of selection. *Animal Genetics, 38*(1), 7-14. doi: 10.1111/j.1365-2052.2006.01543.x

2578

Semina: Ciênc. Agrár. Londrina, v. 43, n. 6, p. 2563-2578, nov./dez. 2022