DOI: 10.1002/vms3.836

ORIGINAL ARTICLE

WILEY

Genetic diversity in reproductive traits of Braunvieh cattle determined with SNP markers

Mitzilin Zuleica Trujano-Chavez[1] | Agustín Ruíz-Flores[1] | Rufino López-Ordaz[1] | Paulino Pérez-Rodríguez[2]

[1]Posgrado en Producción Animal, Universidad Autónoma Chapingo, Texcoco, Estado de México, Mexico

[2]Socio Economía Estadística e Informática, Colegio de Postgraduados, Texcoco, Estado de México, Mexico

Correspondence
Paulino Pérez-Rodríguez, Socio Economía Estadística e Informática, Colegio de Postgraduados, Texcoco, Estado de México, Mexico.
Email: perpdgo@colpos.mx

Agustín Ruíz-Flores, Posgrado en Producción Animal, Universidad Autónoma Chapingo, Texcoco, Estado de México, Mexico.
Email: arf@correo.chapingo.mx

[Correction added on 26 May 2022, after first online publication: Affiliation 2 was updated.]

Funding information
Universidad Autónoma Chapingo, Grant/Award Number: DGIP-166701012

Abstract

Braunvieh is an important dual-purpose breed in the Mexican tropics. The study of its genetic diversity is key to implementing genetic improvement programs. This study was conducted to determine genetic diversity of reproductive traits in a Mexican Braunvieh beef cattle population using single nucleotide polymorphisms in candidate genes. Information from 24 genes with 52 intra-genic loci reported in literature to be associated with productive life, pregnancy rate and cow and heifer conception rate of 150 Braunvieh males and females was considered. Observed heterozygosity (Ho) revealed high genetic diversity for the studied traits, Ho = 0.42 ± 0.087, relative to that of other populations of the same breed. Cluster analyses were carried out using the Ward and K-means algorithms. These analyses revealed high genetic diversity that was observed in the biplot of non-metric multi-dimensional scaling. It was found that clustering strategy allowed visualisation of distant groups by genotype but not by favourable alleles in all the loci. We found that the genes *CSNK1E*, *DNAH11*, *DSC2*, *IBSP* and *OCLN* affected most of the traits in our study and they were highly informative. Therefore, they represent a potential resource for selection and crossbreeding programs of the traits studied in Braunvieh. The analyses showed that the Mexican Braunvieh population has a high level of genetic diversity, arguably due to decades-long adaptation to the Mexican tropics.

KEYWORDS
Brown Swiss, candidate gene, cluster analysis, genetic variability

1 | INTRODUCTION

One of the first objectives in animal breeding was to increase production efficiency per animal (Fleming et al., 2018). Production of animals of high genetic merit allowed improvement of economically important traits, such as milk production in dairy cattle or average daily gain in chickens. However, some traits of similar or higher importance were dismissed. According to Veronese et al. (2019), for years genetic selection in cattle ignored the improvement of reproductive traits.

The strategy in most dairy and dual-purpose herds has emphasised selection of productive and reproductive traits jointly to make the production system more efficient. Since 2000, genetic and genomic selection has included, among others, traits such as cow and heifer conception, obtaining positive results because the highly animals' genetic merit for these reproductive traits (Veronese et al., 2019). Single nucleotide polymorphisms (SNP) have become a widely used resource for genomic selection and have allowed the detection of candidate genes associated with reproductive

traits, as PL1N or ESR2 genes (Raza et al., 2020; Mohammadabadi, 2021).

The study of cattle breeds using a molecular approach is key for their characterising (Mohammadabadi, 2021). Therefore, measuring genetic diversity is essential for genetic improvement, preserving populations, evolution and adapting to variable environments (Olschewsky & Hinrichs 2021; Mohammadabadi et al., 2021). Molecular approaches also allow finding candidate genes associated with traits of economic importance, which could bring a better understanding at the physiological level of the genes' influence. In addition, finding these genes could improve the approach of genetic diversity studies, delineating the research boundaries just for the most important genes (Gooki et al., 2019; Shamsalddini et al., 2016).

Breeds adapted to harsh environments are usually more genetically diverse because they need to be versatile to face changes when they are moved from one environment to another; for instance, 34 Chinese cattle breeds showed high genetic diversity in complex environments (Xia et al., 2018). One of the most demanding production systems for cattle production is the dual-purpose system in tropical regions. The breeds that have had satisfactory results in these conditions are *Bos indicus*, such as Indubrasil (Peixoto et al., 2021) or Guzerat (Campos et al., 2017).

Recently Braunvieh has experienced an important boom in these dual-purpose systems because of its good results for beef production (Rojo-Rubio, 2009). Braunvieh is a breed that is well adapted to cold climates (Bhati et al., 2020). It has been successfully used in dual-purpose systems as the paternal breed (Rojo-Rubio, 2009). Their successful adaptation to tropical conditions suggests that the Mexican population of Braunvieh has important genetic diversity with potential use in genetic improvement programs for reproductive traits in dual-purpose systems.

Genetic diversity studies have been conducted in Braunvieh and its derivative breeds, aiming to unravel the genetic basis of its adaptation to tropical environments. Moscarelli et al. (2020) found important genetic differences among several Braunvieh-derived breeds. Moreover, Bhati et al. (2020) found that Braunvieh has higher genetic diversity than other dual-purpose *Bos taurus* breeds. The objective of our study was to characterise genetically a Mexican Braunvieh population for productive life (PL), pregnancy rate (PR) and cow and heifer conception rate (CCR and HCR), through an analysis of genetic diversity focusing on genotypes and favourable alleles for intra-genic SNPs of candidate genes, using expected and observed heterozygosity indexes, grouping algorithms, analysis of non-metric multi-dimensional scaling (NMDS) and estimation of allele frequencies.

2 | MATERIAL AND METHODS

2.1 | Source of information

Genomic information was obtained from hair samples from 150 animals born between 2001 and 2016 on five farms belonging to the Mexican Association of Braunvieh Purebred Breeders in Eastern, Central and Western Mexico. The samples were genotyped at GeneSeek (Lincoln, NE, USA; http://genomics.neogene.com). The chip used was the Genomic Profile Bovine LD with 50,000 SNP markers.

2.2 | Genotype quality control

The missing genotypes were imputed using the observed allele frequencies. The imputation method consisted of assigning an allele, according to the probability of a possible type of polymorphism for a certain marker in the whole population. Additionally, we considered a 0.05 threshold for the Minor Allele Frequency. For all the genotype quality control, R software version 4.0.4 (R Core Team, 2021) was used.

2.3 | Statistical analyses

2.3.1 | Identification of associated and informative loci

The candidate genes associated with reproductive traits in *Bos taurus* were identified and the position ranked by base pairs were obtained in the Gene Library (Gene, 2021). With this information, the intragenic loci or SNP markers were searched; the criterion for the loci-gene association was that they should be located within the position rank of the gene for the same chromosome. Only the loci available in the genotype information were included. The analyses considered 52 SNP markers of 24 candidate genes reported in the literature (Cochran et al., 2013; Ortega et al. 2015) associated with PR, HCR, CCR and PL.

All the analyses were performed using the R software version 4.0.4 (R Core Team, 2021). The 10 or 12 most informative markers associated, according to the literature consulted, with each of the studied traits were determined. The Shannon index (*H*; Tables 1 and 2) was used. This index allows identification of the most informative markers (Sherwin, 2010). The index is defined as:

$$H = -\sum_{i=0}^{2} p_i \log p_i,$$

where $i = 0, 1, 2$ refers to the genotype ($AA = 0$, $AB = 1$, $BB = 2$), and p_i is the corresponding allelic frequency and log(·) denotes the natural logarithm function. This index was calculated for each of the loci. The 10 or 12 most informative markers were selected. The highest values for H were associated with the highest diversity for the marker.

2.3.2 | Cluster analyses and NMDS

The groups formed by their genetic diversity were generated using hierarchical grouping based on the Euclidian distance matrix. The analysis was performed using the *cluster* function included in the

TABLE 1 Top 12 and 10 single nucleotide polymorphisms (SNP) associated with pregnancy rate and cow productive life in Braunvieh cattle

Pregnancy rate (PR)		Productive life (PL)	
Gene (DA)	SNP	Gene (DA)	SNP
APBB1 (A)	BovineHD1500013531	AP3B1 (A)	BovineHD1000003029
BCAS1 (A)	BovineHD1300023602	DSC2 (B)	BovineHD2400007184
BOLA-DMB (B)	Hapmap60475-rs29022896		BovineHD2400007183
CSNK1E (B)	Hapmap39945-BTA-75021	FSHR (B)	ARS-BFGL-NGS-5623
DNAH11 (A)	BovineHD0400008600		Hapmap45323-BTA-90907
	Hapmap38265-BTA-96973	HSD17B12 (A)	ARS-BFGL-NGS-4967
DSC2° (A)	BovineHD2400007184		BovineHD1500021362
	Hapmap52659-ss46526625	IBSP° (A)	BovineHD0600010277
	BovineHD2400007183	LHCGR (A)	BovineHD1100009276
DZIP3° (A)	BovineHD0100015047	OCLN (A)	BovineHD2000003255
LHCGR° (A)	BovineHD1100009276		
OCLN (B)	BovineHD2000003255		

Gene (DA), gene associated with the reproductive trait, in parenthesis the desirable allele reported by Cochran et al. (2013) and Ortega et al. (2015). Gene°, genes that have been reported with overdominance or dominance.

TABLE 2 Top 10 single nucleotide polymorphisms (SNP) associated with heifer and cow conception rate in Braunvieh cattle

Cow conception rate (CCR)		Heifer conception rate (HCR)	
Gene (DA)	SNP	Gene (DA)	SNP
APBB1° (A)	BovineHD1500013531	CSNK1E (A)	Hapmap39945-BTA-75021
BCAS1 (A)	BovineHD1300023602	DNAH11 (A)	BovineHD0400008600
CSNK1E (A)	Hapmap39945-BTA-75021		Hapmap38265-BTA-96973
DNAH11 (A)	BovineHD0400008600	DSC2° (A)	BovineHD2400007184
	Hapmap38265-BTA-96973		Hapmap52659-ss46526625
DSC2 (B)	BovineHD2400007184	DZIP3° (A)	BovineHD0100015047
	Hapmap52659-ss46526625	FYB1 (B)	BTB-00778141
GOLGA4° (B)	BTB-01184624		BovineHD2000010086
IBSP (A)	BovineHD0600010277	GOLGA4 (B)	BTB-01184624
OCLN (B)	BovineHD2000003255	IBSP (A)	BovineHD0600010277

Gene (DA), gene associated with the reproductive trait, in parenthesis the desirable allele reported by Cochran et al. (2013) and Ortega et al. (2015). Gene°, genes that have been reported with overdominance or dominance.

stats package of R (R Core Team, 2021). Additionally, data were also clustered using the K-means algorithm, using the *K-means* function included in the *stats* package in R. In both cases, the objective was to generate groups of animals with similar genotypic profiles for each of the studied traits. The 10 most informative markers were used to that end.

To validate the results found using hierarchical grouping and the K-means method, the data was also analysed using the Partition Around Medioids (PAM) algorithm (Lengyel and Botta-Dukát, 2019). We performed the analysis using the *pam* function included in the *cluster* package (Maechler et al., 2019) in R. Average values of silhouette width by individual and by group were obtained. The desirable values are close to 1, meaning that on average the objects are close to others in the same group and their classification is correct. On the other hand, values close to −1 indicate that the classification is undoubtedly incorrect (Lengyel and Botta-Dukát, 2019).

NMDS was used for graphic representation in the bi-dimensional space of the distance matrices (per trait), using the *vegan* library in R (Oksanen et al., 2020). The stress value proposed by Kruskal (Dexter et al., 2018) was obtained. This metric estimator indicates how well the algorithm has managed to arrange the points in the ordination while preserving the rank-order distance. The desirable stress values range from 0 to 0.2. Stress values > 0.35 indicate that the samples are randomly placed in the graphic representation and are also considered poor and potentially uninterpretable (Dexter et al., 2018). This multivariant analysis was combined with the K-means results to visualise graphically the clusters and the animal with the representative genotype of each group (the closest to the centroid).

2.3.3 | Heterozygosity and allelic frequencies

Levels of expected (He) and observed (Ho) heterozygosity for the 52 initial markers of the studied population were obtained using the software Cervus 3.0.7 (Kalinowski et al., 2007). Standard deviations were estimated through the non-parametric Bootstrap method for 10,000 replicates using the *boot* function included in the R package *boot* (Canty and Ripley, 2021).

The allelic frequencies for each group obtained with the K-means algorithm were computed for each selected trait using the software Cervus 3.0.7 (Kalinowski et al., 2007). In order to determine whether significant differences existed between allelic frequencies for desirable alleles of the groups formed, and those of representative animals, the X^2 tests for proportions were carried out using the *chisq.test* function included in the *stats* package in R (R Core Team, 2021). The null hypothesis was that non-significant differences exist between proportions of two or more groups (Shih and Fay, 2017). Graphs were generated to visualise proportions for the frequencies of favourable alleles of the most informative genes in the database that affect three to four of the studied traits: *CSNK1E*, *DNAH11*, *DSC2*, *IBSP* and *OCLN*.

2.3.4 | Statement of animal rights

Ethical review and approval were waived for this study because the producers obtained the hair samples for genotyping at their farms, following their procedures. Each of them observed all applicable international, national and/or institutional guidelines for the care and use of animals.

TABLE 3 Genetic diversity indices by marker associated with reproductive traits in the Mexican Braunvieh cattle population: observed (Ho) and expected (He) heterozygosity and their standard deviation SD)

Marker name (locus)	He ± SD	Ho ± SD
ARS-BFGL-NGS-4967	0.466 ± 0.132	0.477 ± 0.100
ARS-BFGL-NGS-5623	0.474 ± 0.135	0.503 ± 0.141
BovineHD0100015047	0.457 ± 0.130	0.420 ± 0.115
BovineHD0400008600	0.497 ± 0.141	0.493 ± 0.110
BovineHD0600010277	0.502 ± 0.144	0.487 ± 0.099
BovineHD1000003029	0.369 ± 0.104	0.313 ± 0.577
BovineHD1100009276	0.490 ± 0.141	0.500 ± 0.140
BovineHD1300023602	0.493 ± 0.141	0.427 ± 0.135
BovineHD1500013531	0.500 ± 0.144	0.497 ± 0.100
BovineHD1500021362	0.457 ± 0.129	0.527 ± 0.142
BovineHD2000003255	0.476 ± 0.100	0.453 ± 0.097
BovineHD2000010086	0.501 ± 0.134	0.520 ± 0.140
BovineHD2400007183	0.452 ± 0.100	0.420 ± 0.105
BovineHD2400007184	0.487 ± 0.101	0.453 ± 0.089
BTB-00778141	0.498 ± 0.110	0.510 ± 0.142
BTB-01184624	0.490 ± 0.098	0.460 ± 0.151
Hapmap38265-BTA-96973	0.499 ± 0.100	0.487 ± 0.100
Hapmap39945-BTA-75021	0.499 ± 0.110	0.500 ± 0.139
Hapmap45323-BTA-90907	0.419 ± 0.100	0.433 ± 0.156
Hapmap52659-ss46526625	0.493 ± 0.123	0.487 ± 0.115
Hapmap60475-rs29022896	0.488 ± 0.121	0.540 ± 0.134
Average for the 52 loci	0.470 ± 0.122	0.420 ± 0.087

3 | RESULTS AND DISCUSSION

3.1 | Identification of associated and informative loci

The 10 most informative loci associated with the studied traits are shown in Tables 1 and 2. For each gene, the favourable allele reported in literature is shown in parentheses. For some of the genes, there are over-dominance reports; therefore, the favourable genotype is the heterozygote. However, genotypes are not transmitted from one generation to the next. Therefore, even though the effect of allele interaction is relevant in crossbreeding breeding systems, our study on genetic diversity focuses on alleles whose additive effect is transmitted from one generation to another.

3.2 | Heterozygosity

Genetic diversity indexes (He and Ho; Table 3) are key parameters in genetic improvement of populations. They were used in our study to determine the level of genetic diversity in the population studied. The genetic diversity measured as heterosis was quite similar among the 21 markers. The frequency of heterozygote individuals in the population was close to 0.5, which indicates that the genetic variation corresponds with what was expected, according to Hardy–Weinberg Equilibrium (HWE). In fact, all the markers in the study at $p > 0.0003$ (Bonferroni correction) were in HWE.

These results suggest that genetic diversity of reproductive traits in the Mexican Braunvieh population is higher than the average of the breeds derived from the original Braunvieh from Switzerland. Moscarelli et al. (2020) reported values close to 0.35 for European populations of original Braunvieh, Braunvieh, Brown Swiss and Italian Brown. This suggests an important difference from our findings, Ho = 0.42 (Table 3). Moscarelli et al. (2020) found that animals of the same breeds that stayed in Switzerland have a higher Ho than those adapted to other environments, such as the Italian Brown, possibly because of conservation of purebred systems. In the population studied the opposite occurred: heterozygosity was higher, possibly due to its adaptation to the tropical environment, which is very different from that in which Braunvieh originated.

The Ho in the Braunvieh population studied for reproductive traits coincides with the findings for other breeds adapted to harsh

environments. Campos et al. (2017) reported Ho for nine *Bos indicus* breeds; their estimates ranged from 0.32 to 0.39, values higher than those found for the breeds related to Braunvieh previously discussed. These results coincide with the findings of our study, although the Ho of the Mexican population is still high. Xia et al. (2018) reported Ho estimates for more than 40 Asian cattle breeds. These authors observed that a value close to 0.5 was associated with breeds, such as Kazakh, Mongolian and Anxi, adapted to the extreme environments of Russia and Mongolia. These authors reported estimates of 0.25 for breeds adapted to temperate climates. In another study, Peixoto et al. (2021) reported values for Ho between 0.47 and 0.85 in Brazilian Guzerat, which is adapted to tropical environments. According to these authors, high genetic variation is typical of Creole populations or those subject to HWE. This could be the reason that Ho for the Braunvieh population of our study is like that of breeds adapted to extreme climates. The Mexican Braunvieh population has had more than a century of adaptation to tropical conditions.

3.3 | Clustering using the Ward and K-means algorithms

The average silhouette width for all the traits obtained with the Ward method was 0.13, whereas it was 0.18 using the K-means method. These estimates suggest that clustering with the K-means algorithm is more adequate than with the Ward method. Although these estimates were low, they were positive in all the cases. Therefore, they can be considered acceptable values that reveal clear genotypic differences between the groups formed with the algorithms (Lengyel and Botta-Dukát, 2019).

Both methods of analysis yielded the same number of clusters, two for PR, HCR and CCR and three for PL. The circular dendrograms obtained from the dissimilarity matrix with the Ward algorithm for the studied traits, allowed grouping the 150 individuals of the study population into different groups (Figure 1).

Clustering algorithms are a useful tool in genetic diversity studies. The findings of the present study coincide with Campos et al. (2017), who used principal components analysis in Zebu cattle in Brazil. They also agree with Öner et al. (2017), who used Unweighted Pair Group Method using Arithmetic Averages to create a phylogenetic tree between Holstein populations, and with Kim et al. (2018), who used the same procedure for Korean breeds.

The results of clustering suggest that the diversity found is sufficient to create at least two groups by trait, which differ in the proportion of animals and in the combination of favourable alleles. In the Figure 1 dendrogram for PR, animal 141 (cluster I) is classified in the smallest group, which contains animals with predominantly BB genotype. In contrast, the largest group was made up of animals in which the AA genotype predominates (Table 4). For PL, three clusters were formed; those containing animals 123 (cluster II) and 133 (cluster III) have the same proportion of genotypes. However, they differ in genotypes for the genes DSC2, HSD17B12, IBSP, LHCGR and OCLN. These two groups differ from cluster I (132 is the representative animal), which does not possess BB genotypes, but only AA and AB genotypes in a 2:3 proportion (Tables 1 and 4). The dendrograms for CCR (d) and HCR (c) show that the groupings are similar. For CCR, the largest group (cluster II), according to the representative animal 112, contains a greater proportion of BB genotypes, whereas cluster I (animal 135) contains a greater proportion of AB genotypes. For HCR, cluster I (animal 54) contains mostly AB genotypes, and AA genotypes are absent. On the other hand, animal 93 is representative of cluster II, which includes mostly AB genotypes.

Table 4 shows the different groups formed with K-means. Given the large number of animals studied, only the five closest to the centroid and the representative genotype of the individual closest to the centroid are reported. In addition, the percentages of desirable genotypes and alleles of each individual are shown to detect the group of animals with mostly favourable genotypes that would be useful in selection.

There are few significant differences in the proportion of desirable alleles and genotypes between the groups formed with the K-means algorithm due to a compensatory effect in the combination of alleles (Moscarelli et al., 2020). In Table 4, it can be observed that the proportion of desirable alleles does not differ significantly (Tables 1 and 2) between groups. An exception was for PL, where cluster III has the lowest significantly different proportion of favourable alleles.

The few differences between proportions are attributed to a compensatory effect, a consequence of analysing the genes as a set to obtain the proportions per group. This procedure was due to the high degree of dissimilarity in genetic variation between markers and genes. Although the main reason for compensation could be that grouping with the K-means method places the animals into different groups because their genotypes are different, generating specific combinations; for example, for PR, cluster I includes a larger number of animals with BB genotypes than cluster II.

These differences do not necessarily mean that any of the groups has a higher number of desirable genotypes or alleles. Alleles vary as a function of the genes; for example, in PR for the gene *BCAS* the desirable allele is A; however, for the *CSNK1E* gene it is B. This means that cluster I contains the favourable allele for *CSNK1E*, but not for *BCAS*, and vice versa for cluster II. The same holds true for most of the cases. That is, the proportions of favourable alleles for all the groups are close to 0.5 (Table 4) due to the compensatory effect determined by the combination of alleles with which this grouping method works (Moscarelli et al., 2020), but it is not related to the number of favourable alleles.

The grouping methods suggest the existence of genetic diversity that would be useful in improvement programs. Despite the compensatory effect previously described, the genetic variation found is good, given that the proportions of favourable genes are important, even though they do not appear assembled together for all the genes. According to Peixoto et al. (2021), the high variation and the eventual formation of groups is an indication of diversity conservation and assures breed sustainability over generations. Today, genetic conservation of adapted or local breeds is a relevant concern due to their contribution to sustainable selection of genotypes.

FIGURE 1 Dendrograms obtained by Ward's algorithm based on Euclidian distance for pregnancy rate (a), productive life (b), cow conception rate (c) and heifer conception rate (d) in a Mexican Braunvieh cattle population

3.4 | Non-metric multi-dimensional scaling

In Figure 2, the bi-dimensional representations obtained with NMDS for the studied traits are shown. To visualise the grouping by genotypes, results from the K-means classification algorithm was also included in the same graphic where we have highlighted each of the groups and the representative animals that belongs to each group (Table 4).

The Kruskal stress values for PR (0.22), CCR and HCR (0.24) are acceptable given they are not greater than 0.35, whereas the value for PL (0.15) permits graphic visualisation of what NMDS offers (Dexter et al., 2018). NMDS is a useful tool to visualise the groups of animals with different genotypes found by the K-means algorithm. Figure 2, from the NMDS analysis, shows that there were no individuals classified in groups different from those assigned by the K-means method. Similarly, the representative animals are close to the centre of the graphic representation of their groups. Although on the borders of the groups some animals are not clearly classified, the groups are well delimited with respect to their neighbouring groups.

NMDS is a tool widely used in genetic diversity studies. Similar results to the ones reported here using NMDS have been reported by other authors, for example, the study of Moscarelli et al. (2020) with breeds derived from the original Braunvieh and the study of Senczuk

TABLE 4 Clustering of 150 individuals (only five representative animals per group are shown) for reproductive traits according to the K-means method, and proportion tests for the desirable genotypes and alleles of representative animals in each group, in a Mexican population of Braunvieh cattle

Trait[1]	Group	NA[2]	Representative animals	Representative genotype	DGP[3]	DAP[4]
PR	I	62	2, 23, 31, 135, 141[⊗]	2 2 2 0 1 0 0 1 1 0 2 1	0.33[a]	0.54[a]
	II	88	1, 24, 50, 74[⊗], 75	1 1 1 0 1 1 1 0 0 1 1 2	0.33[a]	0.63[a]
PL	I	54	21, 33, 48, 50, 132[⊗]	0 0 0 1 1 1 1 1 0 1	0.20[ab]	0.55[ab]
	II	56	26, 35, 62, 104, 123[⊗]	2 2 0 1 2 0 1 1 0 1	0.33[a]	0.65[a]
	III	40	61, 78, 113, 133[⊗], 138	2 1 0 1 2 1 2 0 1 0	0.10[b]	0.45[b]
CCR	I	66	30, 31, 65, 104, 135[⊗]	1 1 0 2 2 1 1 0 1 0	0.20[a]	0.50[a]
	II	84	58, 75, 90, 112[⊗], 134	2 2 1 1 1 2 0 2 0 0	0.33[a]	0.50[a]
HCR	I	96	13, 54[⊗], 75, 91, 129	2 1 1 1 2 1 1 1 1 1	0.20[a]	0.55[a]
	II	54	3, 93[⊗], 104, 141, 146	1 1 2 2 1 1 1 1 0 2	0.20[a]	0.45[a]

[1]PR, pregnancy rate; PL, productive life; CCR = cow conception rate; HCR = heifer conception rate.
[2]NA, number of animals.
[3]DGP, desirable genotype proportion.
[4]DAP, desirable allele proportion. Genotype coding: AA = 0, AB = 1, BB = 2. The order of the loci in representative genotype is the same as in Tables 1 and 2, for each trait. Animal[⊗] is the representative animal (the closest to the centroid). The null hypothesis of non-significant differences between proportions was tested using X^2 and 0.05 significance level; proportions with the same letter are not significantly different.

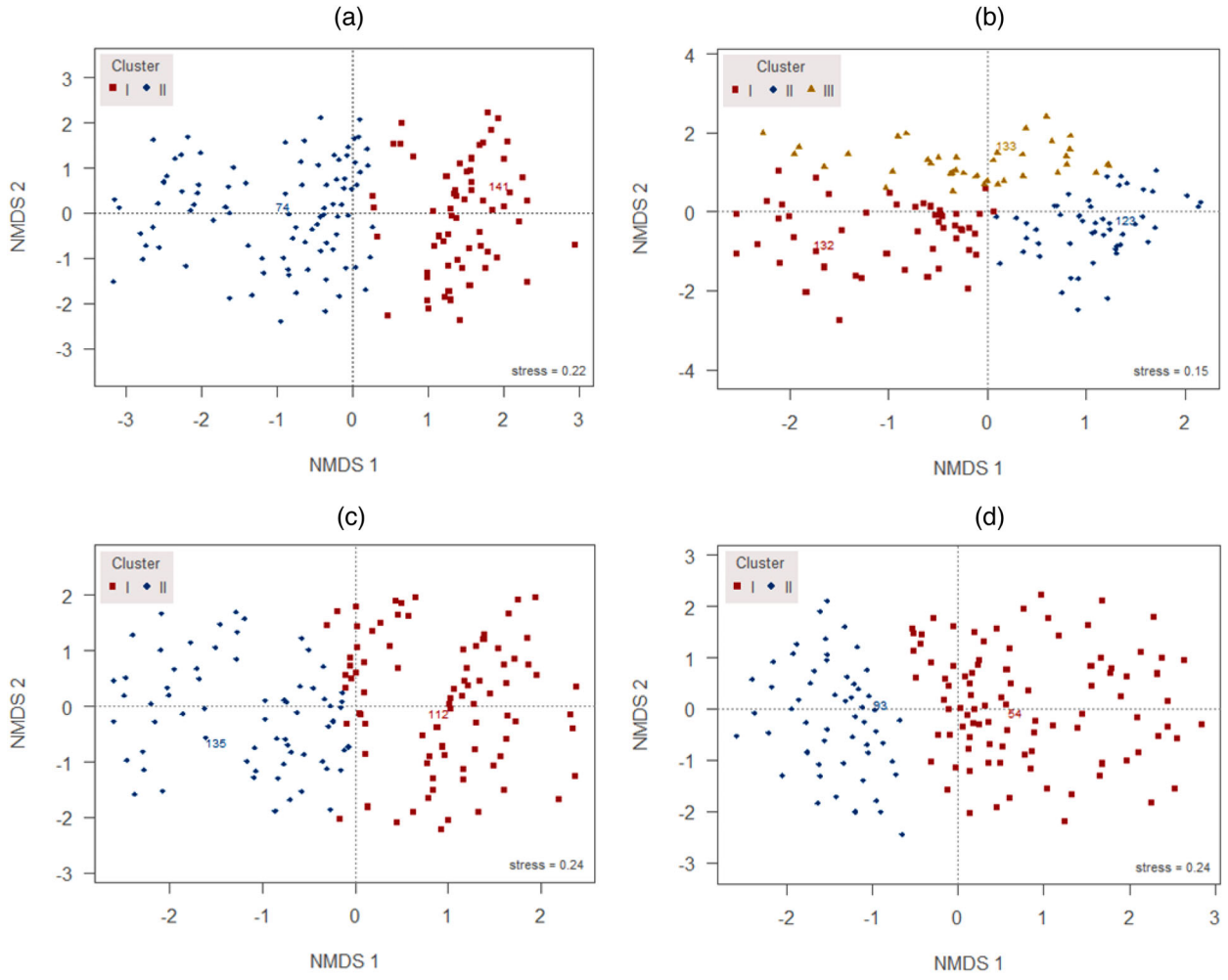

FIGURE 2 Graphic representation of the non-metric multi-dimensional scaling (NMDS) for pregnancy rate (a), productive life (b), cow conception rate (c), and heifer conception rate (d), using K-means clusters to visualise graphic differences and the representative animals, in a Mexican Braunvieh cattle population

TABLE 5 Proportion tests using X^2 test between K-means groups by reproductive trait, for frequencies of desirable alleles per gene in a Mexican Braunvieh population

	PRG		PLG			CCRG		HCRG	
Genes	I	II	I	II	III	I	II	I	II
APBB1	0.42[a]	0.55[a]				0.40[b]	0.61[a]		
AP3B1			0.61[a]	0.51[a]	0.52[a]				
BCAS1	0.42[a]	0.44[a]				0.42[a]	0.45[a]		
BOLA-DMB	0.40[a]	0.59[a]							
CSNK1E	0.43[a]	0.49[a]				0.60[a]	0.46[a]	0.52[a]	0.56[a]
DNAH11	0.51[a]	0.50[a]				0.55[a]	0.45[a]	0.50[a]	0.30[b]
DSC2	0.25[b]	0.63[a]	0.19[c]	0.69[b]	0.88[a]	0.73[a]	0.29[b]	0.65[a]	0.00[b]
DZIP3	0.29[a]	0.39[a]						0.38[a]	0.3[a]
FSHR			0.33[b]	0.62[a]	0.35[b]				
FYB1								0.54[a]	0.45[a]
GOLGA4						0.73[a]	0.38[b]	0.51[b]	0.70[a]
HSD17B12			0.53[a]	0.51[a]	0.31[b]				
IBSP			0.48[a]	0.49[a]	0.33[b]	0.56[a]	0.42[a]	0.48[a]	0.53[a]
LHCGR	0.40[b]	0.61[a]	0.42[c]	0.92[a]	0.59[b]				
OCLN	0.99[a]	0.38[b]	0.32[b]	0.21[b]	0.52[a]	0.52[b]	0.73[a]		
Average	0.46[a]	0.51[a]	0.41[b]	0.57[a]	0.50[a]	0.56[a]	0.47[a]	0.51[a]	0.41[a]

PRG, pregnancy rate groups. PLG, productive life groups. CCRG, cow conception rate groups. HCRG, heifer conception rate groups. The null hypothesis established was that there were non-significant differences between proportions, using X^2 test and 0.05 significance level, proportions with the same letter are not significantly different.

et al. (2020) where NMDS was used to visualise bi-dimensionally differences among cattle breeds from the Alpine arc.

3.5 | Allele frequencies

In Mexico, studies using massive genomic information are scarce. According to Mrode et al. (2019) the main reason is the limited resources available to invest in the genotyping of animals in developing countries. These authors point out that the genetic improvement infrastructure of these countries cannot be compared with the advanced infrastructures of developed countries, whose genomic databases are made up of millions of animals. Given these limitations, the studies of Zepeda-Batista et al. (2019) and Trujano-Chavez et al. (2021) used 300 animals to characterise a Mexican Braunvieh population, the same used in our study, but with a more limited number of animals since the loci found for reproductive traits were only available for 150 animals.

The sample size used in our study was large enough to detect differences between groups. We used the *pwr.p.test* function included in the pwr package (Champely, 2020) to check that the sample size was adequate. The sample size (150) of our study, according to Canals and Canals (2019), was adequate to find significant differences with a 0.05 significance level obtaining a power of the proportion test of 0.92. Other studies like ours found relevant results with even fewer animals. The studies of Agung et al. (2019), Terentjeva et al. (2021) and Xia et al. (2018) had average sample size of 22, 20 and 25 animals, respectively.

Table 5 shows the proportion tests to determine the significant differences between allelic frequencies of desirable alleles per group and gene. Since the genotypes are non-heritable, only allelic frequencies were considered. The genes *DZIP3*, *LHCGR*, *IBSP*, *APBB1*, *GOLGA4* and *DSC2* were reported with dominance and over-dominance effects in *Bos taurus* (Cochran et al., 2013; Ortega et al., 2015). Therefore, the allele considered desirable was that with the greatest additive value for the different traits. Furthermore, the frequencies per gene were considered because the frequencies of the loci or SNP markers within each gene were not significantly different.

Given that the general results (Table 4) for favourable alleles at the genotype level (10 and 12 loci) were non-conclusive because of the compensatory effect, the effect of clustering on changes in allelic frequencies was analysed at the gene/marker level. At the allelic frequency level per gene, there were significant differences ($p < 0.05$) between groups as a product of clustering for PR; group II had the largest number of alleles with favourable effect. The group frequencies for *DSC2*, *LHGR* and *OCLN* were significantly different. That is, the classification of K-means was a function of these three genes. Group II had the largest number of favourable alleles for *DSC2* and *LHCGR*, whereas group I had the highest number for *OCLN*.

Group I had the lowest frequency of favourable alleles for PL and group II the highest. As was previously seen in the first section of grouping, group I for PL had the lowest proportion of favourable alleles when the analysis was performed at the allele combination level (Table 3) and at the gene level (Table 5). In group II, alleles of the genes *FSHR*,

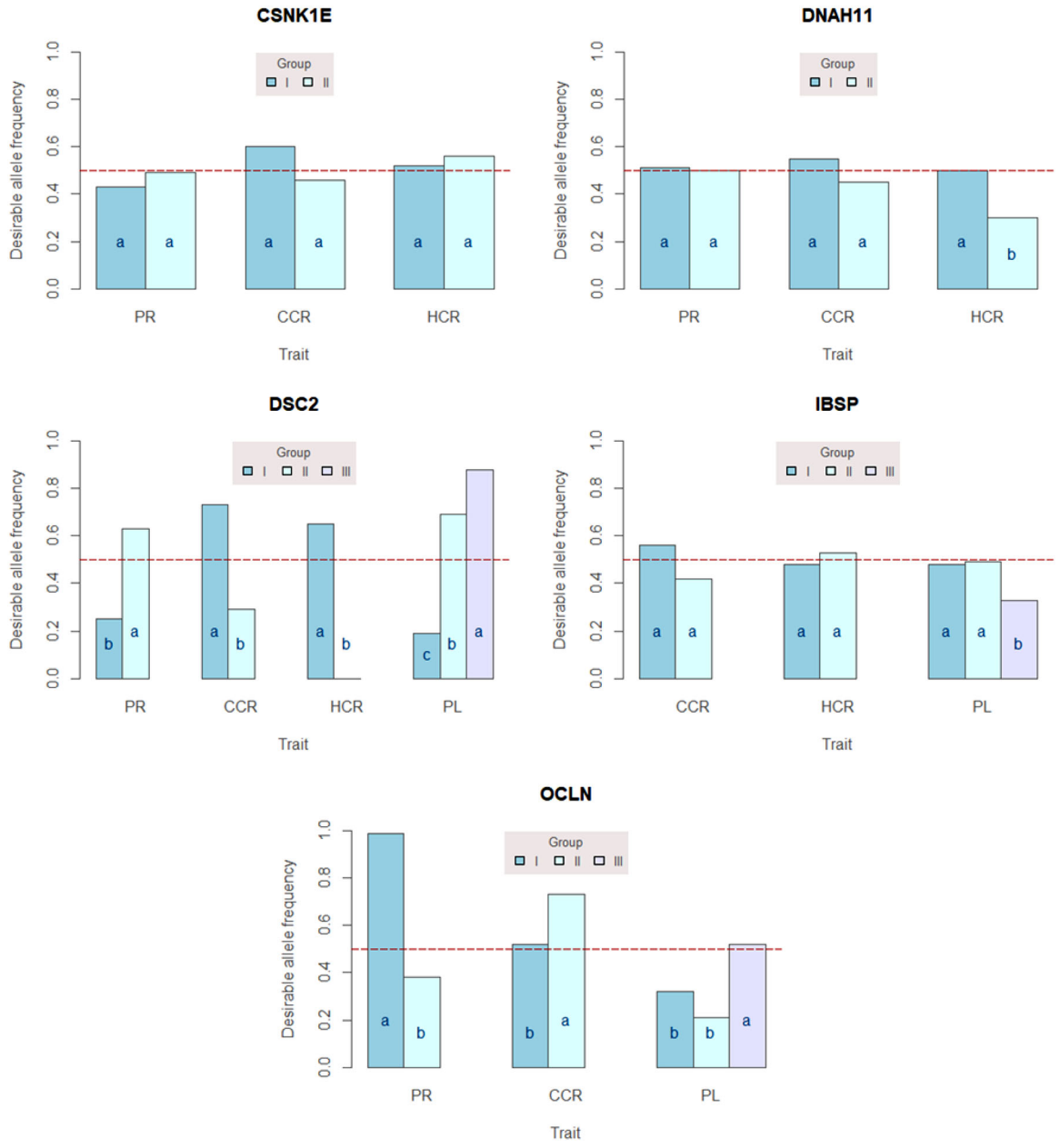

FIGURE 3 Graphic representation of proportion tests using X^2 test between K-means on groups classified by reproductive trait for frequencies of desirable alleles of the most important genes in a Mexican Braunvieh population. PR, pregnancy rate. PL, productive life. CCR, cow conception rate. HCR, heifer conception rate. The null hypothesis established that there were non-significant differences between proportions, using X^2 test and 0.05 of significance level; proportions with the same letter are not significantly different

HSD17B12, *IBSP* and *LHCGR* had higher frequencies. Although group III was better than group I, it had frequencies with significant differences for only *DSC2* and *OCLN*. Therefore, group II would be selected when selection is intended for PL.

The differences between allelic frequencies of the CCR groups are not enough to determine that either of the groups is better. In contrast with the traits previously discussed, the results for CCR are not conclusive by themselves. This is because in both groups two genes exist with significantly higher frequencies (Table 5), and the average for the frequencies is not different ($p > 0.05$) because of the compensatory effect. Therefore, for this trait in particular, the animals should be selected based on the other traits, that is, the animals in favourable groups for the other high-frequency traits for the genes they have in common. For example, if group I were selected for PL, group I animals will be selected indirectly for CCR since *DSC2* is a common gene. For HCR, the results were different from those for CCR. Group I is the selection candidate to improve this trait. Unlike its homologous trait, in HCR group I possesses significant differences for two genes (*DNAH11* and *DSC2*), whereas group II has significant differences only for gene *GOLGA4*.

Some of the genes studied have more relevance for joint selection of animals. Genes *CSNK1E*, *DNAH11*, *DSC2*, *IBSP* and *OCLN* affect two to three of the studied traits. Therefore, selecting exclusively for these

genes could have a generalised impact where selection is for multiple traits. Gene *DSC2*, which affects the four studied traits, could be a direct selection objective to improve reproductive traits. The results of the proportion test for the groups formed by each of the traits for these genes of the higher importance is shown in Figure 3.

Although for the most important genes some of the frequencies of favourable alleles are not significantly different between groups by trait, the results suggest that it is possible to leverage some of the groups of animals by their frequency of favourable alleles. This can be observed for example in Figure 3: for *DSC2* group II for PR, group I for CCR and HCR, and group III for PL. Another gene with differences in frequencies of favourable alleles per group is *OCLN*. Group I for PR could be useful, as is group II for CCR and III for PL.

The genetic diversity found indicates that even though not all the animals have favourable alleles, many of them have very desirable alleles. This information could be enough to start genetic improvement programs for reproductive traits in Braunvieh. The genes considered in this study are the product of selection for variation of, not for addition to, the genetic merit of the animals. In the literature reviewed nearly 24 genes associated with the traits studied were found. However, the Shannon index allowed selection of only the most informative markers of those with greater variation. For this reason, the results of our study may be different from those reported in research with genes of major effects, such as the study of Ortega et al. (2015), where the genes *ACAT2*, *AP3B1* and *OCLN* had a greater impact on the genetic value of Holstein animals for reproductive traits than *DSC2*. Gene *DSC2* is considered the most important of our study because its variation allowed classifying the animals in different groups, and it affects the four studied traits, and not because it had the greatest effect as compared with the other genes.

3.6 | Final considerations

The results of our study suggest the presence of genetic diversity for the studied traits in the population analysed. Genetic diversity studies have shown that there are different Braunvieh lines adapted to different environments. This process has had an important effect on the population variation (Moscarelli et al. 2020; Senczuk et al., 2020). In Mexico, the breed has been slowly introduced into dry and humid tropical areas aiming for its use as a dual-purpose breed. This has modified its allelic combinations, as well the Ho, relative to other Braunvieh populations of the world. Braunvieh is widely in commercial dairy operations because of its milk and beef quality conferred by its Swiss cattle origins, the world's most ancient (Rojo-Rubio et al., 2009). The Ho estimates found in this study suggest adaptation of the Braunvieh population to the Mexican tropics since these values differ from those of its direct ancestors and coincide with those obtained in breeds adapted to extreme conditions. Additionally, the groups formed by the K-means algorithm and the differences among the allele combinations support the existence of genetic variation for reproductive traits, allowing the possibility to select animals for specific reproduction objectives to improve PR, PL, CCR or HCR, or through the gene *DSC2* or any of the representative genes, using group allelic frequencies.

The results of this study constitute a first approach to genetic characterisation for reproductive traits of cattle in Mexico. They are satisfactory because the presence of genetic diversity is considered an opportunity to implement selection and/or crossbreeding plans to improve reproductive performance. Likewise, genetic diversity is essential for conservation of populations of genotypes adapted to different environments. Studies on diversity can provide the information necessary to identify genes associated with traits of interest, in this case, for animals well-adapted to conditions different from those of their origin. Genetic diversity is not a concept associated only with a great number of loci, as in this study where only 52 SNP markers were used. Other studies have been carried out using information of only one gene with few alleles. For example, Takeshima et al. (2018) carried out a study on genetic diversity in Bos indicus cattle in South America using information of only the gene *BoLA-DRB3*.

ACKNOWLEDGEMENTS

Authors thank the Consejo Nacional de Ciencia y Tecnología, México, for the financial support to the first author during her Master of Science studies. The authors also thank the Asociación Mexicana de Criadores de Ganado Suizo de Registro for allowing the use of their databases and the breeders for their kind cooperation in this study. This work was in part supported by a grant from the Universidad Autónoma Chapingo, México (DGIP-166701012).

CONFLICT OF INTEREST

None of the authors has any conflict of interest to declare.

AUTHOR CONTRIBUTIONS

Conceptualisation, *data curation*, *formal analysis*, *investigation*, *methodology*, *resources*, *validation*, *visualisation*, *writing-original draft*: M. Z. T.-C. *Investigation*, *writing-review & editing*: R. L.-O. *Conceptualisation*, *data curation*, *investigation*, *resources*, *project administration*, *writing-review & editing*: A. R.-F. *Data curation*, *validation*, *supervision*, *writing-review & editing*: P. P.-R.

DATA AVAILABILITY STATEMENT

The data are available upon request to the first author.

TRANSPARENT PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1002/vms3.836

ETHICS STATEMENT

All applicable international, national and/or institutional guidelines for the welfare and use of animals were met.

ORCID

Mitzilin Zuleica Trujano-Chavez https://orcid.org/0000-0002-3242-1040
Agustín Ruíz-Flores https://orcid.org/0000-0001-8267-2107
Rufino López-Ordaz https://orcid.org/0000-0001-6077-783X
Paulino Pérez-Rodríguez https://orcid.org/0000-0002-3202-1784

REFERENCES

Agung, P. P., Saputra, F., Zein, M., Wulandari, A. S., Putra, W., Said, S., & Jakaria, J. (2019). Genetic diversity of Indonesian cattle breeds based on microsatellite markers. *Asian-Australasian Journal of Animal Sciences*, *32*(suppl. 4), 467–476. https://doi.org/10.5713/ajas.18.0283

Bhati, M., Kadri, N. K., Crysnanto, D., & Pausch, H. (2020). Assessing genomic diversity and signatures of selection in Original Braunvieh cattle using whole-genome sequencing data. *BMC Genomics*, *21*, 27. https://doi.org/10.1186/s12864-020-6446-y

Campos, B. M., do Carmo, A. S., do Egito, A. A., da Mariante, A. S., do Albuquerque, M. S. M., de Gouveia, S. J. J., Malhado, M. C. H., Verardo, L. L., da Silva, B. M. V. G., & Carneiro, S. P. L. (2017). Genetic diversity, population structure, and correlations between locally adapted zebu and taurine breeds in Brazil using SNP markers. *Tropical Animal Health and Production*, *49*, 1677–1684. https://doi.org/10.1007/s11250-017-1376-7

Canals, C., & Canals, A. (2019). When is n large enough? Looking for the right sample size to estimate proportions. *Journal of Statistical Computation and Simulation*, 1–12. https://doi.org/10.1080/00949655.2019.1602125

Canty, A., & Ripley, B. (2021). boot: Bootstrap R (S-Plus) Functions. R package version 1.3-26.

Champely, S. (2020). pwr: Basic Functions for Power Analysis. R package version 1.3-0. https://CRAN.R-project.org/package=pwr

Cochran, S. D., Cole, J. B., Null, D. J., & Hansen, P. J. (2013). Discovery of single nucleotide polymorphisms in candidate genes associated with fertility and production traits in Holstein cattle. *BMC Genetics*, *14*(suppl. 1), 1–23.

Dexter, E., Rollwagen-Bollens, G., & Bollens, S. M. (2018). The trouble with stress: A flexible method for the evaluation of nonmetric multidimensional scaling. *Limnology and Oceanography: Methods*, *16*(suppl. 7), 434–443. https://doi.org/10.1002/lom3.10257

Fleming, A., Abdalla, E. A., Maltecca, C., & Baes, C. F. (2018). Invited review: Reproductive and genomic technologies to optimize breeding strategies for genetic progress in dairy cattle. *Archives Animal Breeding*, *61*, 43–57. https://doi.org/10.5194/aab-61-43-2018

Gene. (2021). Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information; 2004. Available at: https://www.ncbi.nlm.nih.gov/gene/ Accessed on March 26, 2021

Gooki, F. G., Mohammadabadi, M., Fozi, M. A., & Soflaei, M. (2019). Association of biometric traits with growth hormone gene diversity in Raini Cashmere goats. *Walailak Journal of Science and Technology*, *16*(7), 499–508. https://doi.org/10.48048/wjst.2019.3791

Kalinowski, S. T., Taper, M. L., & Marshall, T. C. (2007). Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology*, *16*, 1099–1106.

Kim, S., Cheong, H. S., Shin, H. D., Lee, S. S., Roh, H. J., Jeon, D. Y., & Cho, C. Y. (2018). Genetic diversity and divergence among Korean cattle breeds assessed using a BovineHD single-nucleotide polymorphism chip. *Asian-Australasian Journal of Animal Sciences*, *31*(suppl. 11), 1691–1699. https://doi.org/10.5713/ajas.17.0419

Lengyel, A., & Botta-Dukát, Z. (2019). Silhouette width using generalized mean—A flexible method for assessing clustering efficiency. *Ecology and Evolution*, *9*(suppl. 23), 13231–13243. https://doi.org/10.1002/ece3.5774

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., & Hornik, K. (2019). cluster: Cluster Analysis Basics and Extensions. R package version 2.1.0.

Mohammadabadi, M. R. (2021). Tissue-specific mRNA expression profile of ESR2 gene in goat. *Agricultural Biotechnology Journal*, *12*, 169–184. https://doi.org/10.22103/JAB.2021.17011.1284

Mohammadabadi, M., Bordbar, F., Jensen, J., Du, M., & Guo, W. (2021). Key genes regulating skeletal muscle development and growth in farm animals. *Animals*, *11*, e835. https://doi.org/10.3390/ani11030835

Moscarelli, A., Sardina, M T., Cassandro, M., Ciani, E., Pilla, F., Senczuck, G., Portolano, B., & Mastrangelo, S. (2020). Genome-wide assessment of diversity and differentiation between original and modern Brown cattle populations. *Animal Genetics*, *52*(1), 21–31. https://doi.org/10.1111/age.13019

Mrode, R., Ojango, J. M. K., Okeyo, A. M., & Mwacharo, J. M. (2019). Genomic selection and use of molecular tools in breeding programs for indigenous and crossbred cattle in developing countries: Current status and future prospects. *Frontiers in Genetics*, *9*, 694. https://doi.org/10.3389/fgene.2018.00694

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Henry, M., Stevens, H., Szoecs, E., & Wagner, H. (2020). Vegan: Community Ecology Package. R package version 2.5-7. https://CRAN.R-project.org/package=vegan

Olschewsky, A., & Hinrichs, D. (2021). An overview of the use of genotyping techniques for assessing genetic diversity in local farm animal breeds. *Animals*, *11*(suppl. 7), 2016. https://doi.org/10.3390/ani11072016

Öner, Y., Keskin, A., Üstüner, H., Soysal, D., & Karakaş, V. (2017). Genetic diversity of the 3' and 5' untranslated regions of the HSP70.1 gene between native Turkish and Holstein Friesian cattle breeds. *South African Journal of Animal Science*, *47*(suppl. 4), 424. https://doi.org/10.4314/sajas.v47i4.2

Ortega, M. S., Denicol, A. C., Cole, J. B., Null, D. J., & Hansen, P. J. (2015). Use of single nucleotide polymorphisms in candidate genes associated with daughter pregnancy rate for prediction of genetic merit for reproduction in Holstein cows. *Animal Genetics*, *47*, 288–297. https://doi.org/10.1111/age.12420

Peixoto, M. G. C. D., Carvalho, M. R. S., Egito, A. A., Steinberg, R. S., Bruneli, F. ÂT., Machado, M. A., Santos, F. C., Rosse, I. C., & Fonseca, P. A. S. (2021). Genetic diversity and population genetic structure of a Guzerá (*Bos indicus*) meta-population. *Animals*, *11*(suppl. 4), 1125. https://doi.org/10.3390/ani11041125

R Core Team. (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Raza, S. H. A., Shijun, L., Khan, R., Schreurs, N. M., Manzari, Z., Abd El-Aziz, A. H., Ullah, I., Kaster, N., Shah, M. A., & Zan, L. (2020). Polymorphism of the PLIN1 gene and its association with body measures and ultrasound carcass traits in Qinchuan beef cattle. *Genome*, *63*(suppl. 10), 483–492. https://doi.org/10.1139/gen-2019-0184

Rojo-Rubio, R., Vázquez-Armijo, J. F., Pérez-Hernández, P., Mendoza-Martínez, G. D., Salem, A. Z. M., Albarrán-Portillo, B., González-Reyna, A., Hernández-Martínez, J., Rebollar-Rebollar, S., Cardoso-Jiménez, D., Dorantes-Coronado, E. J., & Gutierrez-Cedillo, J. G. (2009). Dual purpose cattle production in Mexico. *Tropical Animal Health and Production*, *41*, 715–721. https://doi.org/10.1007/s11250-008-9249-8

Senczuk, G., Mastrangelo, S., Ciani, E., Battaglini, L., Cendron, F., Ciampolini, R., Crepaldi, P., Mantovani, R., Bongioni, G., Pagnacco, G., Portolano, B., Rossoni, A., Pilla, F., & Cassandro, M. (2020). The genetic heritage of Alpine local cattle breeds using genomic SNP data. *Genetics Selection Evolution*, *52*(suppl. 40), 1–12. https://doi.org/10.1186/s12711-020-00559-1

Shamsalddini, S., Mohammadabadi, M. R., & Esmailizadeh, A. K. (2016). Polymorphism of the prolactin gene and its effect on fiber traits in goat. *Russia Journal of Genetics*, *52*, 405–408. https://doi.org/10.7868/s0016675816040093

Sherwin, W. B. (2010). Entropy and information approaches to genetic diversity and its expression: Genomic geography. *Entropy*, *12*(suppl. 7), 1765–1798. https://doi.org/10.3390/e12071765

Shih, J. H., & Fay, M. P. (2017). Pearson's chi-square test and rank correlation inferences for clustered data. *Biometrics*, *73*(suppl. 3), 822–834. https://doi.org/10.1111/biom.12653

Takeshima, S., Corbi-Botto, C., Giovambattista, G., & Aida, Y. (2018). Genetic diversity of BoLA-DRB3 in South American Zebu cattle populations. *BMC Genetics*, *19*(suppl. 13), 1–13. https://doi.org/10.1186/s12863-018-0618-7

Terentjeva, M., Šteingolde, Ž., Meistere, I., Elferts, D., Avsejenko, J., Streikiša, M., Gradovska, S., Alksne, L., Kibilds, J., & Bērziņš, A. (2021). Prevalence, genetic diversity and factors associated with distribution of Listeria monocytogenes and other Listeria spp. in cattle farms in Latvia. *Pathogens*, *10*(suppl. 7), 851. https://doi.org/10.3390/pathogens10070851

Trujano-Chavez, M. Z., Valerio-Hernández, J. E., López-Ordaz, R., Pérez-Rodríguez, P., & Ruíz-Flores, A. (2021). Allelic and genotypic frequencies for loci associated with meat quality in Mexican Braunvieh cattle. *Tropical Animal Health and Production*, *53*(suppl. 2), 307. https://doi.org/10.1007/s11250-021-02757-5

Veronese, A., Marques, A., Moreira, R., Belli, A. L., Bisinotto, R. S., Bilby, T. R., Peñagaricano, F., & Chebel, R. C. (2019). Genomic merit for reproductive traits. I: Estrous characteristics and fertility in Holstein heifers. *Journal of Dairy Science*, *102*(suppl. 7), 6624–6638. https://doi.org/10.3168/jds.2018-15205

Xia, X., Yao, Y., Li, C., Zhang, F., Qu, K., Chen, H., Huang, B., & Lei, C. (2018). Genetic diversity of Chinese cattle revealed by Y-SNP and Y-STR markers. *Animal Genetics*, https://doi.org/10.1111/age.12742

Zepeda-Batista, J. L., Parra-Bracamonte, G. M., Núñez-Domínguez, R., Ramírez-Valverde, R., & Ruíz-Flores, A. (2019). Screening genetic diseases prevalence in Braunvieh cattle. *Tropical Animal Health and Production*, *51*, 25–31. https://doi.org/10.1007/s11250-018-1655-y

How to cite this article: Trujano-Chavez, M. Z., Ruíz-Flores, A., López-Ordaz, R., & Pérez-Rodríguez, P. (2022). Genetic diversity in reproductive traits of Braunvieh cattle determined with SNP markers. *Veterinary Medicine and Science*, *8*, 1709–1720. https://doi.org/10.1002/vms3.836