UNIVERSIDAD AUTÓNOMA CHAPINGO

DEPARTAMENTO DE ENSEÑANZA, INVESTIGACIÓN Y SERVICIO EN ZOOTECNIA

POSGRADO EN PRODUCCIÓN ANIMAL

DIVERSIDAD GENÉTICA PARA CARACTERÍSTICAS DE IMPORTANCIA ECONÓMICA EN BOVINOS SUIZO EUROPEO

TESIS

Que como requisito parcial

para obtener el grado de:

MAESTRO EN CIENCIAS EN INNOVACIÓN GANADERA

Presenta:

MITZILIN ZULEICA TRUJANO CHAVEZ

Bajo la supervisión de: **Agustín Ruíz Flores, Dr.**

y **Paulino Pérez Rodríguez, Dr.**

UNIVERSIDAD AUTÓNOMA CHAPINGO
1854
COORDINACIÓN GENERAL DE ESTUDIOS DE POSGRADO

APROBADA

Chapingo

Chapingo, Estado de México, diciembre 2021

DIVERSIDAD GENÉTICA PARA CARACTERÍSTICAS DE IMPORTANCIA ECONÓMICA EN BOVINOS SUIZO EUROPEO

Tesis realizada por **MITZILIN ZULEICA TRUJANO CHAVEZ**, bajo la supervisión del Comité Asesor indicado, aprobada por el mismo y aceptada como requisito parcial para obtener el grado de:

MAESTRO EN CIENCIAS EN INNOVACIÓN GANADERA

DIRECTOR:

Dr. Agustín Ruíz Flores

CODIRECTOR:

Pérez Rdz

Dr. Paulino Pérez Rodríguez

ASESOR:

Dr. Luis Alberto Miranda Romero

CONTENIDO

LISTA DE CUADROS

LISTA DE FIGURAS

ABREVIATURAS

ADN/DNA	Ácido desoxirribonucleico
	Deoxyribonucleic Acid
AFLP	Polimorfismo en la longitud de fragmentos amplificados de DNA
	Amplified Fragment Length Polymorphism
BC	Corrección de Bonferroni
	Bonferroni Correction
CAPS	Secuencia polimórfica amplificada y cortada
	Cleaved Amplified Polymorphic Sequences
CCR	Tasa de concepción de la vaca
	Cow Conception Rate
F_{IT}	Coeficiente de consanguinidad
	Inbreeding Coefficient
GS	Selección genómica
	Genomic Selection
GWAS	Estudio de asociación del genoma completo
	Genome-Wide Association Study
HCR	Tasa de concepción en vaquillas
	Heifer Conception Rate

He	Heterocigosidad esperada
	Expected heterocigosity
Ho	Heterocigosidad observada
	Observed heterocigosity
HWE	Equilibrio Hardy-Weinberg
	Hardy-Weinberg Equilibrium
MAS	Selección asistida por marcadores
	Marker Assisted Selection
NMDS	Escalamiento multidimensional no métrico
	Non-metric Multidimensional Scaling
PAM	Partición alrededor de medioides
	Partition Around Medioids
PCA	Análisis de componentes principales
	Principal Component Analysis
PL	Vida productiva
	Productive Life
PR	Tasa de preñez
	Pregnancy Rate
QTL	Loci de características cuantitativas
	Quantitative Trait Loci

RAPD	Fragmentos polimórficos de ADN de amplificación aleatoria
	Random Amplification of Polymorphic DNA
RFLP	Polimorfismo de longitud de fragmentos de restricción
	Restriction Fragment Length Polymorphisms
SNP	Polimorfismo de nucleótido simple
	Single Nucleotide Polymorphism
SSR	Repetición de secuencia simple o microsatélite
	Simple sequence repeat
RBM	Resistencia a mastitis bacteriana
	Resistance to Bacterial Mastitis
RCM	Resistencia a mastitis clínica
	Resistance to Clinical Mastitis
RM	Resistencia a mastitis
	Resistance to Mastitis
RSM	Resistencia a mastitis subclínica
	Resistance to Subclinical Mastitis
SM	Susceptibilidad a mastitis
	Susceptibility to Mastitis

DEDICATORIA

A mis padres Enrique Trujano y Teresa Chávez por darme el regalo de la vida, enseñarme a dirigirme con rectitud y buenos valores; además, de brindarme su amor y apoyo siempre incondicionales.

A mi hermana y tía, Violeta Serrano y Marcela Chávez por su apoyo incondicional, amor, enseñanzas y consejos.

A mis sobrinos Naim y Raziel por enseñarme el significado del amor desinteresado y la protección.

A Reyna Sánchez, por ser un pilar indispensable en las nuevas etapas de mi vida, por su apoyo y amor incondicional.

A Diana Olivera, pese a la distancia, la mejor amiga que me ha acompañado en este y otros proyectos de mi vida.

AGRADECIMIENTOS

A la Universidad Autónoma Chapingo, por acogerme y darme más de lo que podía pedir, por permitirme formar parte de su gran comunidad y por brindarme conocimiento.

Al Pueblo de México, que a través del Consejo Nacional de Ciencia y Tecnología (CONACyT), me brindó el apoyo económico para poder realizar mis estudios de Maestría y las presentes investigaciones.

Al Dr. Agustín Ruíz Flores por creer en mí, dándome su apoyo incondicional, supervisión, consejo y motivación en esta investigación y durante toda mi estancia en el Departamento de Zootecnia.

Al Dr. Paulino Pérez Rodríguez por su guía, apoyo, supervisión e indispensable contribución en los análisis estadísticos de esta investigación.

Al Dr. Luis Alberto Miranda Romero por su contribución en esta investigación.

DATOS BIOGRÁFICOS

Mitzilin Zuleica Trujano Chavez

Lugar y fecha de nacimiento: Texcoco, Estado de México; 28 de mayo de 1997.

CURP: TUCM970528MMCRHT07

Profesión: Ingeniero Agrónomo Especialista en Zootecnia

Cédula Profesional: 12203347

Desarrollo Académico

2020-2021 Maestría en Ciencias en Innovación Ganadera. Universidad Autónoma Chapingo.

2012-2019 Ingeniero Agrónomo Especialista en Zootecnia. Universidad Autónoma Chapingo.

Formación Adicional: Cursos

2021 Writing in the Sciences. Stanford University.

2021 The R Programming Environment. Johns Hopkins University.

2021 Matrix Methods. University of Minnesota.

2020 Cálculo Diferencial e Integral unidos por el Teorema Fundamental del Cálculo. Instituto Tecnológico y de Estudios Superiores de Monterrey (ITESM).

2019 Introducción a la Programación Estadística con R. Universidad Nacional Autónoma de México (UNAM).

RESUMEN GENERAL

DIVERSIDAD GENÉTICA PARA CARACTERÍSTICAS DE IMPORTANCIA ECONÓMICA EN BOVINOS SUIZO EUROPEO[1]

El objetivo de la presente investigación fue estudiar la diversidad genética de una población mexicana de bovinos Suizo Europeo para varias características de importancia económica, mediante parámetros poblacionales como heterocigosidad observada, coeficiente de consanguinidad y frecuencia de alelos favorables. Además, se usó estadística multivariada para visualizar la dispersión de los animales con base en su perfil genotípico (análisis de componentes principales, escalamiento multidimensional no métrico y agrupación por los métodos de agrupamiento jerárquico y K-means). Todos los análisis se realizaron con al menos 150 animales genotipados con el chip Genomic Profile Bovine LD de 30K y 50K marcadores de polimorfismo de nucleótido simple. En el primer estudio se encontró que, para muchos de los genes relacionados con calidad de la carne, la población mexicana de Suizo Europeo presentó alta frecuencia alélica favorable para genes asociados con mayor suavidad, mayor marmoleo, menor cantidad de ácidos grasos saturados, y menor fuerza de corte (CAPN1 y DGAT2). En el segundo estudio se midió la diversidad genética para características reproductivas, donde los genes CSNK1E, DNAH11, DSC2, IBSP y OCLN afectaron la mayoría de las características y fueron altamente informativos por lo que se encontraron al menos dos grupos con potencial de selección. En el tercer estudio se determinó la diversidad y estructura genética de la misma población para resistencia y susceptibilidad a mastitis. Si bien, 13.3% de los animales fueron altamente consanguíneos (>50%), hubo grupos altamente heterocigotos, indicativo de diversidad y por tanto favorables. Aunque, los genotipos homocigotos para genes inmunológicos (grupo CXCL) podrían ser favorables para características como resistencia a mastitis clínica y subclínica. En los tres estudios se encontró evidencia del potencial genético de la población mexicana de Suizo Europeo, para mejorar características complejas asociadas con calidad de la carne, reproducción y mastitis.

Palabras clave: consanguinidad, estadística multivariada, genes candidatos, SNP

---

[1] Tesis de Maestría en Ciencias en Innovación Ganadera, Posgrado en Producción Animal, Universidad Autónoma Chapingo
Autor: Mitzilin Zuleica Trujano Chavez
Director: Agustín Ruíz Flores. Codirector: Paulino Pérez Rodríguez

GENERAL ABSTRACT

GENETIC DIVERSITY FOR TRAITS OF ECONOMIC IMPORTANCE IN BRAUNVIEH CATTLE[2]

The aim of this research was to unravel the genetic diversity of a Braunvieh cattle population, using population parameters such as observed heterozygosity, inbreeding coefficient, and favorable allele frequency. Additionally, multivariate statistics were used to visualize the dispersion of the animals based on their genotypic profile using principal component analysis, non-metric multidimensional scaling and grouping by the Brand and K-means methods. All the analyses were performed for at least 150 animals genotyped with the Genomic Profile Bovine LD chip of 30K and 50K single nucleotide polymorphism markers. In the first study it was found that, for many of the genes related to meat quality, the Mexican Braunvieh population presented a high favorable allelic frequency for genes associated with increased tenderness and marbling, less amount of saturated fatty acids, and less cutting force (CAPN1 and DGAT2). In the second study, genetic diversity was determined for reproductive traits, where the CSNK1E, DNAH11, DSC2, IBSP, and OCLN genes affected most of the traits and were highly informative; therefore, at least two groups with selection potential were found. Although 13.3% were highly inbred animals (>50%), there were highly heterozygous groups in terms of the traits studied, a favorable indicator of the presence of genetic diversity. Even though, the homozygous genotypes for immunologic genes (group CXCL) might be favorable for traits such as resistance to subclinical mastitis and resistance to clinical mastitis. In all the three studies, there is evidence of the genetic potential of the Mexican Braunvieh population, to improve complex traits associated with meat quality, reproduction, and mastitis.

Key words: Inbreeding, multivariate statistics, candidate genes, SNP

[2] Master of Science Thesis in Livestock Innovation, Graduate Program in Animal Production, Universidad Autónoma Chapingo
Author: Mitzilin Zuleica Trujano Chavez
Advisor: Agustín Ruíz Flores. Co-advisor: Paulino Pérez Rodríguez

CAPÍTULO 1. INTRODUCCIÓN GENERAL

Durante la historia de selección genética el principal factor para identificar una característica a mejorar ha sido su importancia para generar valor económico. Hasta hace tres décadas en la producción pecuaria se seleccionaban animales sólo para mejorar características productivas como producción de leche, carne, huevo, entre otras. En bovinos Holstein la intensa selección por producción de leche generó la disminución del comportamiento reproductivo de las vacas. Después de la década de 1970, se buscaron otras opciones para mejorar las estrategias de selección para lograr incrementar la ganancia genética de características en forma conjunta (Miglior et al., 2017; Veronese et al., 2019).

De acuerdo con Meuwissen, Hayes y Goddard (2016), algunos de los obstáculos con los que se enfrentó la selección genética en ganado bovino fueron las características cuantitativas con baja heredabilidad, que se miden tarde en la vida del animal y que son difíciles de medir -como las características reproductivas, de calidad de la carne y canal-; por tanto, el mejoramiento basado en fenotipos es un proceso tardado con poca ganancia genética anual.

Desde 1974 con el surgimiento de los primeros marcadores moleculares (Yang et al., 2013), empezó a desarrollarse una nueva estrategia para mejorar características complejas. El uso de los marcadores permitió el mejoramiento genético de animales basándose en su genotipo real, más que en la información fenotípica, lo que atendió en parte el problema planteado por Meuwissen et al. (2016). Además, la asociación de marcadores moleculares con características de importancia económica (donde se obtienen genes candidatos) permite desentrañar la base genética de los procesos fisiológicos relacionados con las características (Yan et al., 2013).

Actualmente, el genotipado de animales es un procedimiento sencillo, basta con tomar una muestra de pelo o sangre de los animales para obtener sus genotipos (Meuwissen et al., 2016). Los marcadores más utilizados para selección genómica son los polimorfismos de nucleótido simple (SNP), dado que su uso

generalizado ha reducido hasta un nivel rentable los costos de genotipado en sistemas comerciales y de investigación (Wiggans et al., 2017).

De acuerdo con Ali et al. (2019), el aprovechamiento de la variación genética es un elemento clave en el desarrollo y uso de cualquier raza en programas de mejoramiento. Para determinar los niveles de variación se han realizado múltiples estudios de diversidad y estructura genética usando la información de los SNP en ganado bovino. Estos estudios se desarrollaron principalmente para características complejas como reproducción, salud, reducción de emisiones de gases de efecto invernadero y resistencia a parásitos (Weller et al., 2017; Xia et al., 2018).

Las razas adaptadas a ambientes adversos poseen mayor variabilidad genética (Xia et al., 2018), debido a la necesidad del organismo para ser versátil ante los cambios del ambiente. En México, uno de los sistemas que afronta mayores adversidades climáticas para la producción bovina es el de doble propósito, ubicado en las regiones tropicales del país. Los bovinos adaptados a estos ambientes con resultados satisfactorios son *Bos indicus*, como la raza Indubrasil (Peixoto et al., 2021) o la Guzerat (Campos et al., 2017). Sin embargo, actualmente la raza Suizo Europeo ha tenido un auge importante en estos sistemas de producción debido a los buenos resultados en producción de carne como raza paterna (Rojo-Rubio, 2009).

La raza Suizo Europeo, a pesar de estar adaptada a climas templado-fríos (Bhati et al., 2020), se ha consolidado como una de las favoritas en los sistemas doble propósito como raza paterna (Rojo-Rubio, 2009). Por tanto, se esperaría que la población mexicana de Suizo Europeo en estudio tuviera una variabilidad genética importante, que pudiera aprovecharse en programas de mejoramiento.

Utilizando ejemplares de la raza Suizo Europeo y sus razas emparentadas, se hicieron estudios de diversidad genética para desentrañar la base genética de su adaptación a los ambientes. En el estudio de Moscarelli et al. (2020), se encontraron diferencias genéticas importantes entre varias razas derivadas de la

Braunvieh original, de acuerdo con el ambiente donde terminaron de adaptarse. Por otro lado, Bhati et al. (2020), en un estudio de diversidad con 49 toros ancestros, encontraron que la raza posee alta diversidad genética, frente a otras razas europeas de doble propósito.

Objetivo general

Determinar la diversidad genética de una población de bovinos Suizo Europeo, mediante parámetros poblacionales como heterocigosidad observada, coeficiente de consanguinidad y frecuencia de alelos favorable; además, de usar estadística multivariada para visualizar la dispersión de los animales con base en su perfil genotípico.

Objetivos específicos

1. Estimar las frecuencias génicas y genotípicas para loci de genes candidatos de calidad de la carne en una población mexicana de bovinos Suizo Europeo.
2. Caracterizar genéticamente una población mexicana de Suizo Europeo para vida productiva, tasa de preñez, y la tasa de concepción en vacas y vaquillas, mediante un estudio de diversidad genética enfocado en genotipos y alelos favorables para SNPs intragénicos (±25K pb) de genes candidatos, usando los índices de diversidad de heterocigosidad esperada y observada, algoritmos de agrupamiento, análisis de escalamiento multidimensional no métrico y estimación de frecuencias génicas.
3. Caracterizar genéticamente a la población mexicana de Suizo Europeo para las características resistencia y susceptibilidad a mastitis, resistencia a mastitis bacteriana y mastitis subclínica, y mastitis clínica para SNPs intragénicos (±25K pb) de genes candidatos, usando los índices de diversidad de heterocigosidad esperada y observada, coeficientes de consanguinidad, algoritmos de agrupamiento y análisis de componentes principales.

Este documento de tesis está compuesto por seis capítulos. El Capítulo 2 contiene la revisión de literatura sobre marcadores moleculares, estudios de asociación del genoma completo y el desarrollo de genes candidatos; además, se revisa de manera general los principales descubrimientos de genes candidatos para el grupo de características en estudio; este capítulo concluye con una breve descripción sobre los tipos de estudios de diversidad realizados en el mundo.

En el Capítulo 3 se describe la caracterización genética a través de las frecuencias de alelos favorables para Suizo Europeo. El Capítulo 4 caracteriza a esta población de Suizo Europeo para características reproductivas de acuerdo con sus alelos favorables, se reportan algunos genes con mayor oportunidad de selección por la diversidad de sus alelos.

El Capítulo 5 describe la diversidad y estructura genética de la población bovina estudiada para resistencia y susceptibilidad a mastitis, que a diferencia de los dos capítulos anteriores tiene un enfoque conservativo de versatilidad asociada con los genes más heterocigóticos. Finalmente, en el Capítulo 6 se presentan las conclusiones generales de los estudios que componen esta tesis (capítulos 3 a 5).

Literatura Citada

Ali, I., Suhail, S. M., & Shafiq, M. (2019). Heritability estimates and genetic correlations of various production and reproductive traits of different grades of dairy cattle reared under subtropical condition. *Reproduction in Domestic Animals*, 54, 1026-1033. doi:10.1111/rda.13458

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. doi: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., … & 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. doi:10.1007/s11250-017-1376-7

Meuwissen, T. H. E., Hayes, B. J., and Goddard M. E. (2016). Genomic selection: A paradigm shift in animal breeding. *Animal Frontiers*, 6, 6-14. doi:10.2527/af.2016-0002

Miglior, F., Fleming, A., Malchiodi, F., Brito, L. F., Martin, P., & Baes, C. F. (2017). A 100-Year Review: Identification and genetic selection of economically important traits in dairy cattle. *Journal of Dairy Science*, 100(12), 10251–10271. doi:10.3168/jds.2017-12968

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

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

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

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

Weller, J. I., Ezra, E., & Ron, M. (2017). Invited review: A perspective on the future of genomic selection in dairy cattle. *Journal of Dairy Science*, 100(11), 8633–8644. doi:10.3168/jds.2017-12879

Wiggans, G. R., Cole, J. B., Hubbard, S. M., & Sonstegard, T. S. (2017). Genomic selection in dairy cattle: the USDA experience. *Annual Review of Animal Biosciences*, 5, 309-327.

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*. doi:10.1111/age.12742

Yang, W., Kang, X., Yang, Q., Lin, Y., & Fang, M. (2013). Review on the development of genotyping methods for assessing farm animal diversity. *Journal of Animal Science and Biotechnology*, 4(1), 2. doi:10.1186/2049-1891-4-2

CAPÍTULO 2. REVISIÓN DE LITERATURA

2.1 Marcadores moleculares

De acuerdo con Singh et al. (2014), un marcador genético es un término general aplicable a un fenotipo con base genética, que ayuda a evaluar la variabilidad fenotípica observada en los individuos. Los mismos autores señalan que existen diferentes tipos de marcadores genéticos que son base del mejoramiento genético animal, estos pueden ser morfológicos, productivos, bioquímicos y moleculares.

Un marcador molecular es producto del avance en metodologías moleculares para análisis genético, actualmente es lo más preciso para explicar los fenotipos de los animales (Marwal & Gaur, 2020). En concreto, un marcador molecular es una variación en un segmento de ADN -o gen con ubicación conocida- específica entre individuos asociada con características de importancia económica (Nadeem et al., 2017; Singh et al., 2014).

Singh et al. (2014), señalan que estas variaciones en el ADN pueden ser deleciones, inserciones, translocaciones, duplicaciones y mutaciones puntuales. En función de estas variaciones existen diferentes tipos de marcadores, cuyas propiedades varían; en general la historia de los marcadores se puede resumir en tres grupos, que se describen a continuación.

2.1.1 Marcadores RFLP, RAPD y AFLP

Según Yang, Kang, Yang & Fanq (2013), dentro de los marcadores obtenidos mediante hibridación, se encuentran los de polimorfismo de longitud de fragmentos de restricción (RFLP; por sus siglas en inglés), que fueron los primeros marcadores basados en ADN y se desarrollaron en 1974. De acuerdo con Singh et al. (2014), estos marcadores se basan en inserciones o deleciones en los sitios de restricción de las enzimas, donde se utilizan sondas para la identificación de los genes y la eventual genotipificación.

Las principales ventajas de los marcadores RFLP: son altamente reproducibles, codominantes, multialélicos y no requieren información de secuencia para generarse; mientras que, algunas de sus desventajas son: se requiere gran cantidad de ADN de alta calidad, sus procesos de obtención son tardados y difíciles de automatizar (Hashim & Al-Shuhaib, 2019; Yang et al., 2013).

Yang et al. (2013) mencionan que en 1990 se desarrollaron por primera vez los marcadores de fragmentos polimórficos de ADN amplificados aleatoriamente (RAPD; por sus siglas en inglés), que se basan en la amplificación del ADN mediante PCR, usando un primer único de 10 nucleótidos. Nadeem et al. (2017) señalan que, una ventaja importante que comparten estos marcadores con los RFLP es que no necesitan información de secuencias para su desarrollo; además, el proceso de obtención de los RPD no involucra la técnica de la hibridación, es decir, es rápido, simple y eficiente. Sin embargo, poseen una desventaja importante, es un marcador dominante, lo que limita su aplicación en el mejoramiento animal a características explicadas por esta acción génica.

Un tercer marcador llegó a minimizar las desventajas en la obtención y aplicación de los RFLP y RAPD. Los marcadores de polimorfismo en la longitud de fragmentos amplificados de ADN (AFLP; por sus siglas en inglés), desarrollados en 1993, son la combinación de polimorfismo en sitios de restricción e hibridación de *primers* arbitrarios, es decir, en los AFLP se combina el poder de los RFLP con la simplicidad en el proceso de obtención de los AFPL (Nadeem et al., 2017). A pesar de estas ventajas, los AFLP son marcadores dominantes lo que limita su uso generalizado (Nadeem et al., 2017).

Yang et al. (2013) mencionan que los AFLPs destacan por su estabilidad genética, su desarrollo rápido y barato; además, a pesar de ser estrictamente dominantes, los AFLPs pueden ser útiles en genética de poblaciones y la caracterización del genoma completo, para evaluar y caracterizar los recursos genéticos animales. A pesar de estas aplicaciones, los AFLPs actualmente han

caído en desuso debido al surgimiento de otros marcadores de ADN que brindan información genética más precisa y de forma eficaz.

2.1.2 Marcadores SSR o microsatélites

Ante las limitantes de los marcadores ya mencionados, surgió un nuevo tipo de marcador, llamado microsatélite o repeticiones de secuencia simple (SSR; por sus siglas en inglés). Estos marcadores son repeticiones cortas en tándem conformadas por, repeticiones de 2 a 3 bases (dinucleótidos y trinucleótidos) comúnmente. Los SSR se caracterizan por ser altamente polimórficos (hasta más de 20 alelos), debido a la gran cantidad de repeticiones en el genoma (Nadeem et al., 2017).

Los SSR poseen las mismas ventajas que los RFLP, tienen mayor repetibilidad y estabilidad que los RAPD, además de ser marcadores codominantes. Algunas de sus desventajas son: alto índice de error en el genotipado ante mutaciones no identificadas y desarrollo lento y costoso. Los microsatélites fueron los marcadores más utilizados durante años, hasta la llegada de los polimorfismos de nucleótido simple (SNP; por sus siglas en inglés) (Nadeem et al., 2017; Yang et al., 2013).

2.1.3 Marcadores SNP

Los marcadores SNP son lo último en precisión a nivel de ADN, que representan el cambio de un solo par de bases presentes en la secuencia del genoma de un individuo; desarrollados por primera vez en 1996, son la unidad más pequeña de herencia, y su variación se puede deber a transiciones o transversiones sobre la base de sustitución de nucleótidos. Los métodos de genotipificación SNP más conocidos son: mediante desarrollo de RFLP (SNP-RFLP) y la técnica del marcador de secuencia polimórfica amplificada y cortada (CAPS; por sus siglas en inglés). El método consiste en secuenciar a los animales y encontrar diferencias en un solo par de bases entre ellos (Marwal & Gaur, 2020; Nadeem et al., 2017, Yang et al., 2013;).

De acuerdo con Yang et al. (2013), los marcadores SNP actualmente son de los marcadores más utilizados en mejoramiento genético, dado que son abundantes en el genoma, son genéticamente estables y cuyo análisis puede ser completamente automatizado, por ende, su desarrollo es de alto rendimiento; además, son marcadores codominantes por lo que es sencillo identificar animales heterocigóticos y homocigóticos.

2.2 Estudios de asociación del genoma completo

De acuerdo con Tian et al. (2020), el estudio de asociación del genoma completo (GWAS; por sus siglas en inglés) es una técnica genética clave para explorar la base genética de características de interés en los animales. Estos mismos autores señalan que GWAS pretende detectar la asociación de variaciones genéticas específicas con un fenotipo, el método implica el análisis de genomas de muchos individuos para realizar la búsqueda de marcadores moleculares (sobre todo SNPs), útiles para predecir un valor fenotípico asociado con una característica en estudio.

2.2.1 Metodología básica de GWAS

Uitterlinden (2016) describe un procedimiento básico para realizar un GWAS, el cual involucra los siguientes pasos:

1. *Colección del ADN/genotipado:* este autor señala que actualmente existen muchas opciones para realizar el genotipado, en todas las especies y con varios tipos de muestra disponibles, usualmente se recurre a las muestras de pelo y sangre. Existen diversos chips para obtener los genotipos, entre los que destacan Affymetrix e Illumina. Para encontrar resultados reveladores en un GWAS es clave tener genotipos de dos grupos de individuos: un grupo control y un grupo con la característica deseada.
2. *Imputación de genotipos:* en muchas ocasiones la información genómica resultante de la amplificación de muestras de ADN no es

concluyente, lo que sugiere el desconocimiento de los genotipos para ciertos marcadores en algunos individuos. Imputar los genotipos faltantes es una alternativa viable para mejorar la precisión de GWAS (Uitterlinden, 2016); de acuerdo con Browning et al. (2018), Beagle 5.2 es uno de los programas más usados para imputar genotipos.

3. *Análisis de datos:* hasta este punto, se tienen dos bases de datos, la primera, contiene la información genotípica de los individuos; mientras que, la segunda, contiene el valor fenotípico de la característica de interés para esos mismos individuos. La información resultante del genotipado debe recodificarse para pasar de una escala categórica (AA, AB, BB) a valores numéricos (0, 1 y 2) en la que se cuenta la presencia de cierto alelo.

De acuerdo con la naturaleza de los fenotipos, es posible realizar regresión logística (valores dicotómicos) o regresión lineal (valores continuos); cualquier programa de análisis estadístico es útil mientras pueda procesar grandes matrices. Típicamente hay dos formas de representar los resultados de un GWAS: representados como valores de p para la prueba de asociación (observados vs. esperados) en una gráfica cuantil-cuantil o como una gráfica de Manhattan (Uitterlinden, 2016).

Un último paso no enumerado por Uitterlinden (2016), es la caracterización del SNP significativo para la característica. Una vez que se ha identificado que un SNP posee un efecto sobre la característica estudiada, se procede a estudiar este segmento de ADN. Recordando que ese marcador específicamente puede ser la causa de variación fenotípica, o bien algún otro segmento de ADN localizado cerca de él, gracias al desequilibrio de ligamiento. Este último paso requiere de múltiples estudios, debido a que es posible que el segmento tenga carácter cuantitativo, lo que involucra mayor trabajo de laboratorio para encontrar los procedimientos fisiológicos que le dan lugar.

2.2.2 Métodos estadísticos para GWAS

En el subtema anterior se comentó acerca del análisis estadístico para realizar GWAS. La regresión lineal simple es el método más elemental para realizar la asociación de los marcadores con los fenotipos. Al respecto Hayes (2013), describe el modelo estadístico para la regresión de marcador único, cuyo uso se fundamenta en la explotación del desequilibrio de ligamiento.

Hayes (2013) señala que, el desequilibrio de ligamiento se refiere a las asociaciones en una población entre marcadores moleculares y mutaciones causales o loci de características cuantitativas (QTL; por sus siglas en inglés). Tales asociaciones surgen debido a que algunos segmentos del genoma -provenientes de un mismo cromosoma- se heredan juntos a través de las generaciones, es decir no hay efectos recombinantes. A estos segmentos de cromosoma se les conoce como haplotipos, usados en el modelo de GWAS más sencillo: regresión marcador por marcador.

Esta regresión, que estudia la asociación entre marcadores moleculares y fenotipos, está dada por (Hayes, 2013):

$$\boldsymbol{y} = \boldsymbol{W}\boldsymbol{b} + \boldsymbol{X}\boldsymbol{g} + \boldsymbol{e},$$

donde $\boldsymbol{y}$ es el vector de fenotipos, $\boldsymbol{W}$ es la matriz de diseño que relaciona los datos con los efectos fijos fenotípicos, $\boldsymbol{b}$ es un vector de efectos fijos como la media, efectos de estructura poblacional y edad, $\boldsymbol{X}$ es el vector que conecta los datos con los efectos de marcador, g es el efecto del marcador y $\boldsymbol{e}$ es el vector de desviaciones aleatorias con el supuesto $e_i \sim N(0, \sigma_e^2)$, donde σ_e^2 es la varianza del error.

La hipótesis nula para el modelo establece que el marcador no tiene efecto en la característica; mientras que, la hipótesis alterna establece un efecto significativo sobre la característica, lo anterior se afirma debido a la existencia de desequilibrio de ligamiento entre los marcadores y un posible QTL (Hayes, 2013).

El modelo anterior es un análisis de un solo SNP por vez, en consecuencia, el enfoque de GWAS generalmente usa SNPs individuales para evaluar su asociación con un fenotipo de interés mientras ignora otros SNPs. Sin embargo, la mayoría de las características de interés en animales son cuantitativas, es decir, una parte no conocida de la heredabilidad incluye efectos epistáticos además de aditivos. Otra limitante del uso de este modelo GWAS es la gran capacidad computacional necesaria para llevar a cabo las operaciones matriciales del modelo mixto (Mortezaei & Tavallaei, 2021).

Para solucionar estos problemas se han desarrollado metodologías que permiten dar un enfoque multi-loci a GWAS y que ayudan a minimizar el problema de grandes bases de datos en la obtención de los resultados. Al respecto, Mortezaei y Tavallaei (2021) enlistan varios tipos de procedimientos alternativos para la realización de GWAS, dos de los cuales se muestran a continuación.

2.2.2.1 Aprendizaje automático

Usando el aprendizaje automático es posible descubrir patrones ocultos dentro de la base de datos de los marcadores, que pueden ayudar a desentrañar la base genética completa de las características. El aprendizaje automático se basa en la inteligencia artificial, donde las computadoras pueden tomar decisiones aprendiendo de los datos. La regresión de marcador único es una forma incipiente del aprendizaje automático; mientras que bosques aleatorios o metodologías basadas en el principio de máxima verosimilitud son herramientas nuevas con aplicaciones novedosas. Las máquina de soporte vectorial son modelos que pueden usarse en conjunto con la regresión simple usada en marcador único (Díez et al., 2021; Mortezaei & Tavallaei, 2021).

2.2.2.2 Análisis de asociación retrospectivo

Considerando los múltiples efectos no genéticos asociados con un fenotipo, se han creado modelos de asociación que incluyen covariables para corregir posibles variaciones no controladas en el ambiente, a estos procedimientos se

les conoce como análisis de asociación retrospectivo, que prometen ser más precisos que los modelos comunes prospectivos para variables binarias; sin embargo, esto no se ha comprobado. Dos de las metodologías estadísticas para realizar este tipo de análisis son la prueba de características binarias longitudinales basada en ecuaciones y la prueba de asociación basada en modelos mixtos lineales generalizados retrospectivos (Mortezaei & Tavallaei, 2021).

2.2.3 Genes candidatos

Derivado de GWAS, los genes candidatos son aquellos que se asocian con una característica. Los genes son segmentos de ADN con ubicación fija en el genoma, se pueden ubicar mediante el número de cromosoma y su rango de posición en pares de bases. En los estudios de asociación, los marcadores usados se encuentran dispersos casi siempre de forma aleatoria a lo largo del genoma, cuando se encuentra un marcador significativo para una característica, se recupera la posición de este marcador ± 25,000 pares de bases y entonces se asigna a un gen que lo contenga. Este gen que contiene el marcador significativo se conoce como gen candidato. Los genes candidatos contienen por lo menos un segmento de ADN que afecta una característica, gracias al desequilibrio de ligamiento (Mota et al., 2017a).

Es más efectivo estudiar genes candidatos que marcadores por sí solos. Existen genes que codifican productos funcionales que tienen efecto sobre diferentes características en el organismo (Gerstein et al., 2007). Por ejemplo, el gen CAPN1, es un gen candidato en bovinos para suavidad y marmoleo de la carne (Page et al., 2004). Fisiológicamente podrían existir hasta dos explicaciones diferentes sobre el efecto del gen y seguramente estas características poseen una correlación positiva dado que los segmentos de ADN que las define están en el mismo gen y poseen equilibrio de ligamiento.

2.3 Genes candidatos en bovinos

Muchos de los genes que se abordarán a continuación, no fueron incluidos en los estudios de los capítulos 3 a 5, debido a que no se encontraban sus posiciones en la base de datos de la que se dispuso para el estudio. Con ello se trata de presentar una perspectiva más completa de lo encontrado para *Bos taurus* referente a los genes candidatos más importantes por característica.

2.3.1 Calidad de la carne

Page et al. (2004) y White et al. (2005), identificaron en el ganado *Bos taurus* y *Bos indicus* SNP en la posición del gen CAPN que se asocian con suavidad de la carne, y estos SNP se les llamó CAPN1-316 y CAPN1-4751. Por su parte, Soria et al. (2010) reportaron otro SNP, asociado al mismo fenotipo, el SNP CAPN1-530.

El alelo C, tanto en CAPN1-316 (Page et al., 2004) como en CAPN1-4751, se ha asociado con valores favorables de terneza de la carne. Sin embargo, este alelo se ha encontrado en baja proporción en algunas poblaciones (Casas et al., 2006; Page et al., 2004; White et al., 2005).

Para el gen CAST se identificó un SNP de sustitución de citosina por timina que está asociado con la terneza de la carne. Los animales homocigotos para timina producen carne con menores valores de fuerza de corte que aquellos homocigotos para citosina (Casas et al., 2006).

El sistema CAPN/CAST está involucrado en la miogénesis y la comunicación entre células, la diferenciación celular y la degradación de proteínas musculares y de proteínas de la membrana celular durante la fusión de los mioblastos, al formarse el miotubo, y de células satélites (Moyen et al., 2004, Barnoy et al., 2005). Según lo reportado por Casas et al. (2006), los polimorfismos relacionados con este sistema pueden influir en el índice de producción y suavidad de la carne.

Por su parte, el gen DGAT1 codifica la proteína acetil Coenzima A diacilglicerol aciltransferasa (DGAT) que cataliza el paso final en la síntesis de triacilglicerol (Casos et al., 1998; Grisart et al., 2002). De acuerdo con Casos et al. (1998), esta enzima también puede estar involucrada en la síntesis y almacenamiento de lípidos, y dos marcadores de este gen se han asociado con la variación en el contenido de grasa de la leche y la carne. Al respecto, Curi et al. (2011) encontraron que el gen DGAT1 es polimórfico en ganado Nelore y que tiene una posible asociación con la cantidad de grasa de la carne.

Los polimorfismos del gen TCAP se encuentran entre los factores genéticos importantes involucrados en la calidad de la canal en el ganado vacuno; el gen titin-cap codifica telethonin, una proteína que interactúa con la titina en la línea Z; las asociaciones entre los SNP del gen TCAP y el veteado en *Bos taurus* mostraron que la sustitución de guanina por adenina redujo la puntuación de veteado (Cheong et al., 2007).

La familia MYOD, conformada por los genes MYOD1 o MYF3, MYOG, MYF4 (miogenina), MYF5 y MYF6 (herculina), juegan un papel clave en el crecimiento y el desarrollo muscular, se consideran genes candidatos para las características de producción de carne en animales (Bhuiyan et al., 2009). La familia está involucrada en la regulación de la hipertrofia muscular y la hiperplasia (Bhuiyan et al., 2009; Houba & Te Pas, 2004).

Según Houba y Te Pas, (2004), el gen MYOG se expresa antes del nacimiento; mientras que, Te Pas y Soumillion (2001) afirman que, en el período posnatal, puede asociarse con la reparación del daño en las fibras musculares y el crecimiento hipertrófico.

En cerdos, MYOG se ha asociado con un aumento en el número y el área transversal de las fibras musculares (Kim et al., 2009), el peso al nacer y en canal, y la tasa de crecimiento (Te Pas et al., 1999). Sin embargo, en el ganado *Bos taurus*, el efecto se observó solo en la capacidad de retención de agua y la suavidad de la carne (Ujan et al., 2011).

Las anquirinas son proteínas involucradas en las interacciones entre la membrana plasmática y el citoesqueleto (Bagnato et al., 2003). Según Rubtsov y Lopina (2000), los genes ANK1, ANK2 y ANK3 codifican anquirinas expresadas en tejidos específicos. El empalme alternativo de ANK1 produce ankyrin que tiene un dominio sensible a la proteólisis y puede ser el objetivo de la calpaína, lo que afecta la suavidad de la carne en bovinos.

De acuerdo con Curi et al. (2011), el gen FABP4 codifica la proteína de unión a ácidos grasos que se encuentra en los adipocitos, sus funciones incluyen la absorción, el transporte y el metabolismo de los ácidos grasos. Curi et al. (2011) reportaron que en ganado Nelore puro el gen puede tener un efecto notorio; aunque el I polimorfismo FABP4/Msp A1I no mostró asociación con las características de área del ojo de la costilla, cantidad de grasa, grasa intramuscular y fuerza de corte e índice de fragmentación miofibrilar en animales cruzados.

2.3.2 Características reproductivas

Entre las características reproductivas de mayor importancia en bovinos están: tasa de concepción para vaquillas y vacas adultas, vida productiva y mérito neto global (Cochran, Cole, Null & Hansen, 2013). Este tipo de características poseen baja heredabilidad y toma mucho tiempo medirlas durante la vida del individuo, la precisión en la predicción también es baja, por lo que encontrar genes candidatos fue la primera opción para mejorar las predicciones y el mejoramiento genético, sobre todo en la raza Holstein (Cochran et al., 2013).

En estudios donde se utiliza GWAS, se han probado miles de genes candidatos reportados en literatura, encontrando de 20 a 50 genes candidatos relacionados de forma importante con características reproductivas (Cochran et al., 2013; Pimentel et al., 2011).

Para tasa de concepción, Ortega et al. (2016) evaluaron diferentes marcadores SNP de genes candidatos en una población de vacas Holstein. Cerca de 57% de

los SNP tuvieron un efecto importante en la característica. Los genes candidatos con efecto significativo fueron: COQ9, EPAS1, CAST, C7H19orf60 y MRPL48. El gen CAST se había discutido en la sección de genes candidatos para calidad de la carne, lo que refuerza la idea antes planteada de que los genes candidatos son un factor de estudio importante, que pueden afectar varias características cuantitativas.

Para la característica de tasa de concepción en vacas, en el mismo estudio de Ortega et al. (2016), se confirmó un efecto importante de los genes candidatos GOLGA4, COQ9, EPAS1 y MRLP48 en la característica. Estos mismos autores reportan genes candidatos para la tasa de concepción en vaquillas HSD17B7, AP3B1, HSD17B12 y CACNA1D. Algunos de los genes reportados asociados con fertilidad en general, están involucrados en la esteroidogénesis o están regulados por esteroides, lo que explica que algunos de estos genes coincidan con genes candidatos asociados con características sobre contenido de grasa (Ortega et al., 2016).

En ganado Canchim -raza compuesta por 62.5% Charolais y 37.5% Cebú- Buzanskas et al. (2017) encontraron que el gen RAP1B participa en varias vías que afectan el comportamiento reproductivo, la principal vía es la del receptor de la hormona liberadora de gonadotropinas. Por otro lado, los mismos autores señalan que MED30 y TRHR también son genes candidatos importantes relacionados con la hormona tiroidea que afecta directamente la reproducción.

Buzanskas et al. (2017), al igual que Ortega et al. (2016), hallaron que los genes FABP5, FABP12, PEX2 y MED30 están involucrados en el metabolismo de ácidos grasos, colesterol y triacilglicerol para la producción de hormonas sexuales, sobre todo para la producción de testosterona.

Según Buzanskas et al. (2017), el gen que afecta significativamente el desempeño reproductivo de machos Canchim es UBQLN3, que se expresa en los testículos y actúa en la espermatogénesis. De acuerdo con estos autores,

este gen está involucrado en el procesamiento de proteínas en la vía del retículo endoplásmico.

2.3.3 Resistencia y susceptibilidad a mastitis

En ganado bovino la mastitis es una enfermedad importante de la glándula mamaria, que causa pérdidas económicas y cuya base genética empezó a estudiarse hace relativamente poco. Algunos estudios han revelado genes candidatos (GC, NPFFR2, TRAPPC, ARHGAP39, LY6K, LY6D, LYNX1, LYPD2, SLURP1 y PSCA) asociados con el conteo de células somáticas para ganado Holstein (Cai, Guldbrandtsen, Lund, & Sahana, 2018; Moretti et al., 2021).

Moretti et al. (2021), encontraron genes candidatos asociados con el conteo en células somáticas en ganado Holstein. Los SNP se identificaron en regiones intragénicas de genes relacionados con componentes clave del sistema inmunológico: CXCR1, DCK, NOD2, MBL2, MBL1 y M-SAA3.2. Por ejemplo, los receptores de quimiocinas CXCR 1 y 2 son genes parálogos que codifican los principales receptores de citocinas proinflamatorias (Moretti et al., 2021), que impactan directamente el proceso de inflamación de la glándula mamaria.

Kurz et al. (2019), encontraron grupos de animales Holstein resistentes y susceptibles a mastitis clínica midiendo la longitud del pezón como variable de respuesta. Encontrando que 7.6% de las vacas presentaban un genotipo para resistencia y 12.5% presentaron susceptibilidad genética para la enfermedad. Estos mismos autores encontraron que el gen mayormente asociado con la longitud del pezón, RASGRP1, es un gen candidato de nuevo descubrimiento que podría ser clave para seleccionar animales con genotipos favorables para disminuir la tasa de incidencia de mastitis en los hatos.

La mastitis es una enfermedad multifactorial, y por ende difícil de desentrañar genéticamente. Algunos enfoques que se han tomado para atender esta problemática se relacionan con la resistencia a bacterias específicas que promueven la aparición de la mastitis. Por ejemplo, Liabin, Chen y Chen (2019)

identificaron genes candidatos asociados con mastitis provocada por *Escherichia coli*, encontrando genes distintos a los que regularmente se hallan en los estudios de GWAS para resistencia en general. Los genes candidatos encontrados fueron IL8RB, CXL6y y MMP9.

Un estudio de Song et al. (2016) reveló genes candidatos asociados con la susceptibilidad a mastitis subclínica provocada por *Staphylococcus aureus* en ganado Holstein. Se encontró que los genes NRG1, MST1 y NAT9 se correlacionaron de manera importante con la característica.

2.3.4 Resistencia a garrapatas

Otra característica difícil de mejorar mediante mejoramiento genético convencional es la resistencia y susceptibilidad a garrapatas. Es difícil de medir y su heredabilidad es baja. Con las metodologías actuales de asociación del genoma e identificación de genes candidatos, el mejoramiento de este tipo de características se hace cada vez más factible. Al igual que en el caso de mastitis, los enfoques se han hecho más específicos para poder dar solución a los diferentes impactos de las garrapatas (Moré et al., 2019).

Los genes asociados con resistencia a garrapatas hallados por Moré et al. (2019) en un estudio con bovinos Braford, tienen influencia en el sistema inmunológico en los animales. En particular, los autores mencionan una vía de señalización de *Wingless* como un mecanismo candidato a resistencia, donde está involucrado el gen WNT7A y los genes IL6 e IL22 como promotores.

En el estudio de Mota et al. (2017b) se encontraron genes candidatos asociados con la resistencia y susceptibilidad a garrapatas en ganado Hereford y Braford, algunos de los genes asociados con resistencia más importantes fueron: CCL20, MYO6, BSP1, BSP3 y PLAUR; mientras que, los asociados con susceptibilidad fueron DNAH8, BOLA-DQA2 y SALL4.

2.4 Uso de los marcadores moleculares en mejoramiento genético animal

2.4.1 Selección asistida por marcadores

Los marcadores moleculares con el surgimiento de la selección asistida por marcadores (MAS; por sus siglas en inglés), prometieron en sus inicios ayudar de forma efectiva y rápida al mejoramiento genético para características importantes economicamente en animales y plantas (Jena & Mackill, 2008). De acuerdo con Jena y Mackill (2008), el propósito en MAS es monitorear la presencia o ausencia de genes de interés en poblaciones; por otro lado, Collard y Mackill (2008) definieron MAS como el proceso de usar marcadores de ADN para asistir en la selección de material para el mejoramiento genético.

En párrafos anteriores se discutieron algunos elementos esenciales para realizar MAS: el surgimiento de los marcadores moleculares, el mapeo de QTL y la identificación de genes candidatos. MAS es una metodología más eficiente que la evaluación fenotípica convencional, que conlleva un ahorro de tiempo, esfuerzo, recursos y dinero (Platten, Cobb & Zantua, 2019).

El primer elemento es el tipo de marcador a utilizar, actualmente el uso de SSR y SNP es lo más pertinente debido a que poseen características deseables para MAS, son confiables, necesitan poca cantidad y calidad de ADN, su procedimiento técnico es sencillo, poseen un alto nivel de polimorfismo y su costo los vuelve rentables (Collard & Mackill, 2008).

Otro elemento importante para el éxito de MAS es el mapeo de QTLs. Este procedimiento se realiza gracias al mapeo de ligamiento que se basa en la recombinación ocurrida durante la meiosis. Un mapeo de QTLs es un procedimiento que permite encontrar loci asociados con características de interés; por lo general este mapeo no es suficiente para designar a un segmento de ADN candidato definitivo, sino que se necesita realizar mapeo fino para ajustarse a un segmento de ADN aún más pequeño (Collard & Mackill, 2008; Platten et al., 2019).

Collard & Mackill (2008) describieron un flujo general para concretar el proceso de MAS:

1. *Desarrollo de las poblaciones:* selección parental e hibridación. Los estudios de MAS pretenden obtener la mayor cantidad de individuos heterocigóticos por lo que se busca obtenerlos mediante selección y cruza.
2. *Mapeo de QTLs:* construcción del mapa de ligamiento, evaluación fenotípica para la característica y análisis de QTL. Con este paso se desea obtener un primer acercamiento de los efectos del QTL con la característica.
3. *Validación de QTLs:* confirmación de la posición y efecto de los QTL, mapeo fino. En este paso se obtiene de forma más precisa la longitud y posición de los QTLs con efectos sobre la característica. Esta validación permite obtener resultados similares entre poblaciones similares.
4. *Validación de marcadores:* aprovechando el desequilibrio de ligamiento y teniendo las posiciones de los QTLs, es posible obtener marcadores que estén asociados con estos QTL y por tanto, tengan un efecto en la característica.
5. *Selección asistida por marcadores*: la información de marcadores es útil en la selección de individuos con genotipos favorables.

Actualmente MAS puede ser una metodología enriquecida haciendo uso de los genes candidatos, es decir de GWAS. Para características difíciles de mejorar mediante el método convencional, como las que se discutieron previamente, MAS puede ser de ayuda para seleccionar individuos con genotipos favorables en etapas tempranas, logrando ahorrar tiempo y dinero en el mejoramiento genético. MAS es una alternativa viable en el mejoramiento de plantas y animales (Collard & Mackill, 2008; Platten et al., 2019). Sin embargo, MAS tiene una desventaja importante, muchas de las características que trata de explicar son cuantitativas, es decir, están influenciadas por muchos genes, lo más grave es que un solo gen no llega a explicar más de 10% de la variación de una característica (Collard & Mackill, 2008).

2.4.2 Selección genómica

La selección genómica (GS; por sus siglas en inglés) se refiere al uso de marcadores que cubren el genoma de los animales para predecir el valor como reproductores de los candidatos a selección de una característica cuantitativa, asumiendo que todos los QTL están en desequilibrio de ligamiento con al menos un marcador. La GS es posible gracias al gran número de SNPs que cubren los genomas de especies domésticas de interés, y los métodos existentes para determinar los genotipos eficientemente en un número grande de marcadores (Goddard, 2009; Goddard & Hayes, 2007).

Como una solución a las limitantes de MAS, la GS permite el mejoramiento de los animales usando toda la información disponible del genotipado. Para incluir la información de algunos marcadores en las predicciones, es posible localizar mutaciones causantes subyacentes a la variación genética o mapeo de QTLs. En la práctica estos programas de mejoramiento genético han tenido un impacto menor al esperado, lo que se atribuye a que la proporción de la varianza contabilizada por el QTL mapeado generalmente es pequeña y que los recursos financieros para realizar los mapeos son elevados, lo que limita el uso de MAS (Berry et al., 2016; de los Campos et al., 2013).

En GS se asume que todos los marcadores están unidos a un gen con efectos para una característica, concentrándose en la estimación de estos efectos más que en probar su significancia. La mayoría de las características se ven afectadas por un gran número de genes con efectos pequeños; la predicción de características complejas requiere considerar un gran número de variantes al mismo tiempo (de los Campos et al., 2013; Meuwissen et al., 2016).

La selección genómica según Meuwissen et al. (2016) se hace en dos etapas, la primera incluye el genotipado y fenotipado de animales de la población de entrenamiento para una determinada característica con el fin de estimar los efectos de los SNP. En la segunda etapa se genotipa a los individuos candidatos a selección, población de evaluación, al combinar la información obtenida en la

primera etapa (efectos SNP) y los genotipos se obtienen los valores genómicos que servirán para seleccionar los mejores individuos para la característica evaluada.

Las principales ventajas de GS son la reducción del intervalo entre generaciones, el incremento de la tasa de progreso genético, mayor exactitud en la predicción de valores genéticos, mejora en la exactitud de valores genómicos de animales jóvenes, especialmente para características de heredabilidad baja, limitadas por el sexo, medidas tarde en la vida de los animales, costosas de medir, o que requieren retar a los animales (Garrick, 2011; Garrick & Saatchi, 2011; Yin & König, 2016).

Según Meuwissen et al. (2016), existen tres condiciones que han permitido el amplio uso de información del ADN en las evaluaciones genómicas: 1) la metodología para realizar GS (Meuwissen et al., 2001), 2) la identificación de miles de marcadores SNP, y 3) las tecnologías de costo bajo para determinar eficientemente genotipos.

2.5 Estudios de diversidad genética en bovinos

La diversidad genética se refiere a los diferentes polimorfismos presentes en los genes de individuos de una misma especie, o incluso de una raza o una población. En la actualidad los estudios de diversidad genética se han generalizado debido al surgimiento de marcadores moleculares de bajo costo y con buena distribución en el genoma para especies domésticas (SSR y SNP) (Doublet et al., 2019).

Existen dos enfoques principales para el estudio de la diversidad genética usando marcadores moleculares en bovinos. Un enfoque se relaciona con la consanguinidad y estructura de la población, los estudios se enfocan en observar cómo ha cambiado el coeficiente de consanguinidad en una población a lo largo de los años, o después de implementar un programa de mejoramiento (Doublet et al., 2019; Lozada-Soto et al., 2021). Por otro lado, existe el enfoque

conservacionista que pretende investigar la diversidad genética de razas relacionadas genéticamente cuyos ancestros comunes datan de hace cientos o miles de años, lo que pretende desentrañar la base genética de su adaptación a diferentes regiones (Bhati, Kadri, Crysnanto, & Pausch, 2020; Freitas et al., 2021; Moscarelli et al., 2020).

En el caso de la presente serie de investigaciones, los parámetros relacionados con diversidad y estructura genética se utilizaron para observar la tendencia de los genotipos de marcadores SNP para genes candidatos asociados con características de importancia económica, encontrando valores para heterocigosidad observada (Ho) similares a los de animales adaptados al trópico, aun cuando Suizo Europeo es una raza Bos taurus (Trujano-Chavez et al., 2021). Este enfoque es similar al utilizado por Freitas et al. (2021), quienes estudiaron la diversidad genética de razas adaptadas a diferentes ambientes para la característica de resistencia a estrés calórico.

La diversidad genética por lo regular se auxilia de parámetros poblacionales y visuales para poder ser caracterizada. Depende del enfoque que se desee tomar, es el tipo de parámetros a utilizar y las herramientas visuales propias de la estadística multivariada pertinentes. Generalmente, el coeficiente de consanguinidad y la Ho son dos parámetros que no pueden faltar en un estudio de diversidad genética; mientras que, el análisis de componentes principales (PCA, por sus siglas en inglés) y los árboles filogenéticos son los principales recursos gráficos usados para visualizar las relaciones genéticas (Bhati et al., 2020; Freitas et al., 2021).

2.5.1 Estadísticas básicas de la población

En un estudio de diversidad genética (enfoque de conservación) con 10 razas bovinas y usando microsatélites, Agung et al. (2019) usaron los siguientes parámetros poblacionales: número observado de alelos, número efectivo de alelos, Ho, heterocigosidad esperada (He), equilibrio Hardy-Weinberg. El número observado y efectivo de alelos son parámetros recurrentes en los estudios que

usan microsatélites, los cuáles poseen altos grados de polimorfismos, por lo que el conteo de alelos es necesario para describir la diversidad genética. Por otro lado, Ho, He y el equilibrio Hardy-Weinberg son parámetros necesarios para cualquier tipo de estudio de diversidad, inclusive es posible calcularlos sin poseer información de marcadores moleculares.

Un enfoque similar al del párrafo anterior, fue propuesto por Freitas et al. (2021), quienes usaron la Ho, He y el equilibrio Hardy-Weinberg como medidas de la diversidad genética en ganado europeo usando SNPs. Usaron otros parámetros interesantes como la distancia genética media por pares, el desequilibrio de ligamiento y el tamaño efectivo de la población. Lo anterior, para caracterizar la diversidad genética para el estrés térmico en tres razas.

Lozada-Soto et al. (2021) usando un enfoque de evaluación de estructura poblacional, caracterizaron la diversidad genética de ganado Angus usando los coeficientes de consanguinidad basados en pedigrí y en marcadores SNP; además, midieron los intervalos generacionales, el tamaño efectivo de población y el número efectivo de segmentos cromosómicos segregantes independientes. Estos autores se focalizaron en determinar los efectos de la GS sobre la diversidad genética de bovinos Angus, basándose en los coeficientes de consanguinidad más que en la frecuencia en heterocigosidad en los individuos.

En contraste, Xia et al. (2019), usaron el genoma mitocondrial para caracterizar a la raza Yunling, y los parámetros que utilizaron fueron menos convencionales que los de estudios mencionados en párrafos anteriores. Crearon haplotipos a partir de la secuenciación del genoma mitocondrial, para poder medir el número de haplotipos y los sitios variables, la diversidad de nucleótidos y el número medio de diferencias de nucleótidos.

2.5.2 Programas y estadística multivariable

En la red existen una diversidad de *software* para calcular de forma sencilla los principales parámetros poblacionales usados en genética. Estos programas

también ofrecen herramientas estadísticas visuales para facilitar la interpretación de las relaciones genéticas entre los individuos en estudio.

El Boom de R (R Core Team, 2021) ha hecho que muchos paquetes incluyan en sus análisis herramientas para estudiar la diversidad genética. Uno de estos programas es *adegenet*, que ofrece el cálculo de las estadísticas poblacionales más sencillas como coeficientes de consanguinidad, Ho, He y el equilibrio Hardy-Weinberg. Es versátil y ofrece estos parámetros para estudios que usan tanto SNPs como SSRs. Para estos últimos también calcula el número observado y efectivo de alelos (Jombart, 2008).

El paquete *adegenet* en principio fue creado para el análisis multivariado de marcadores, por lo que una de sus principales fortalezas es proporcionar herramientas visuales para la estructura y diversidad de la población. Lo hace ampliando el paquete ade4 y lo vuelve capaz de analizar marcadores moleculares.

Otro paquete de R de amplio uso en los estudios de diversidad genética es *diveRsity*, que promete ser un paquete útil en la estimación y exploración de parámetros de genética de poblaciones y sus errores asociados (Keenan et al., 2013). En esencia, es un programa que calcula los mismos parámetros que *adegenet*; sin embargo, incorpora información para la validación estadística de sus resultados, de acuerdo con Keenan et al. (2013), calcula intervalos de confianza *Bootstrap* por locus y a nivel global.

Además de R, existen otros paquetes útiles en estudios de diversidad como el programa CERVUS, útil en el cálculo de frecuencias génica, Ho, He, equilibrio Hardy-Weinberg y el coeficiente de consanguinidad. POPGEN ofrece cálculos similares, además de calcular el número observado y efectivo de alelos, y el flujo de genes. Los programas MEGA y STRUCTURE son útiles en la visualización gráfica, producen dendogramas y generan grupos de animales de acuerdo con su perfil genotípico usando estadística multivariada (Agung et al., 2019).

Otro programa usado es PLINK, líder en estudios de genética, posee desde análisis especializados como GWAS hasta el cálculo de parámetros básicos como lo hace *adegenet*. PLINK es capaz de realizar cualquier método multivariado para la visualización de los datos, PCA, dendogramas, árboles filogenéticos, escalamientos multidimensionales y aprendizaje automático (Freitas et al., 2021).

2.6 Literatura Citada

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(4), 467–476. doi:10.5713/ajas.18.0283

Bagnato, P., Barone, V., Giacomello, E., Rossi, D., & Sorrentino, V. (2003). Binding of an ankyrin-1 isoform to obscurin suggests a molecular link between the sarcoplasmic reticulum and myofibrils in striated muscles. *Journal of Cell Biology*, 160, 245-253. doi:10.1083/jcb.200208109

Barnoy, S., Maki, M., & Kosower, N. S. (2005). Overexpression of calpastatin inhibits L8 myoblast fusion. *Biochemical and Biophysical Research Communications*, 332(3), 697-701. doi:10.1016/j.bbrc.2005.05.010.

Berry, D. P., Garcia, J. F., & Garrick, D. J. (2016). Development and implementation of genomic predictions in beef cattle. *Animal Frontiers*, 6(1), 32–38. doi:10.2527/af.2016-0005

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. doi:10.1186/s12864-020-6446-y

Bhuiyan, M. S. A., Yu, S. L., Jeon, J. T., Yoon, D., Cho, Y. M., Park, E. W., … & Lee, J. H. (2009). DNA Polymorphisms in SREBF1 and FASN Genes Affect Fatty Acid Composition in Korean Cattle (Hanwoo). *Asian-Australasian Journal of Animal Sciences*, 22(6), 765-773. doi: 10.5713/ajas.2009.80573

Browning, B. L., Zhou, Y., & Browning, S.R. (2018). A one-penny imputed genome from next generation reference panels. *The American Journal of Human Genetics*, 103(3), 338-348. doi:10.1016/j.ajhg.2018.07.015

Buzanskas, M. E., Grossi, D. do A., Ventura, R. V., Schenkel, F. S., Chud, T. C. S., Stafuzza, N. B., … & Munari, D. P. (2017). Candidate genes for male and female reproductive traits in Canchim beef cattle. *Journal of Animal Science and Biotechnology*, 8, 67. doi:10.1186/s40104-017-0199-8

Cai, Z., Guldbrandtsen, B., Lund, M. S., & Sahana, G. (2018). Prioritizing candidate genes post-GWAS using multiple sources of data for mastitis resistance in dairy cattle. *BMC Genomics*, 19, 656. doi:10.1186/s12864-018-5050-x

Casas, E., White, S. N., Wheeler, T. L., Shackelford, S. D., Koohmaraie, M., Riley, D. G., … & Smith, T. P. L. (2006). Effects of calpastatin and calpain markers in beef cattle on tenderness traits. *Journal of Animal Science*, 84(3):520-525. doi:10.2527/2006.843520x

Casos, S., Smith, S. J., Zheng, Y. W., Myers, H. M., Lear, S. R., Sande, E., …& Farese Jr., R. V. (1998). Identification of a gene encoding an acyl CoA: diacylglycerol acyltransferase, a key enzyme in triacylglycerol synthesis.

Proceedings of the National Academy of Sciences of the United States of America, 95, 13018-13023.

Cheong, H. S., Yoon, D., Kim, L. H., Park, B. L., Lee, H. W., Han, C. S., … & Shin, H. D. (2007). Titin-cap (TCAP) polymorphisms associated with marbling score of beef. *Meat Science*, 77(2), 257–263. doi:10.1016/j.meatsci.2007.03.014

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(1), 49-72. doi:10.1186/1471-2156-14-49

Collard, B. C. Y., & Mackill, D. J. (2008). Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. *Philosophical Transactions of the Royal Society B*: *Biological Sciences*, 363(1491), 557–572. doi:10.1098/rstb.2007.2170

Curi, R. A., Chardulo, L. A. L., Arrigoni, M. D. B., Silveira, A. C. & de Oliveira, H. N. (2011). Associations between LEP, DGAT1 and FABP4 gene polymorphisms and carcass and meat traits in Nelore and crossbred beef cattle. *Livestock Science*, 135, 244-250.

de Los Campos, G., Hickey, J. M, Pong-Wong, R., Daetwyler, H. D., & Calus, M. P. (2013). Whole-genome regression and prediction methods applied to plant and animal breeding. *Genetics*, 193(2), 327-345. doi:10.1534/genetics.112.143313

Díez, D. F., Sánchez L. F., Moreno, V., Moratalla-Navarro, F., Molina de la Torre, A. J., & Martín Sánchez, V. (2021). GASVeM: A new machine learning methodology for multi-SNP analysis of GWAS data based on genetic algorithms and support vector machines. *Mathematics*, 9(6), 654. doi:10.3390/math9060654

Doublet, A.-C., Croiseau, P., Fritz, S., Michenet, A., Hozé, C., Danchin-Burge, C., … & Restoux, G. (2019). The impact of genomic selection on genetic diversity and genetic gain in three French dairy cattle breeds. *Genetics Selection Evolution*, 51, 52. doi:10.1186/s12711-019-0495-1

Freitas, P. H. F., Wang, Y., Yan, P., Oliveira, H. R., Schenkel, F. S., Zhang, Y., … Brito, L. F. (2021). Genetic diversity and signatures of selection for thermal stress in cattle and other two Bos species adapted to divergent climatic conditions. *Frontiers in Genetics*, 12, 102. doi:10.3389/fgene.2021.604823

Garrick, D. J. (2011). The nature, scope and impact of genomic prediction in beef cattle in the United States. *Genetics Selection Evolution*, 43, 17-27. doi:10.1186/1297-9686-43-17

Garrick, D. J., & Saatchi, M. (2011). Opportunities and challenges for genomic selection of beef cattle. *Revista Brasileira de Zootecnia*, 40, 310-316.

Gerstein, M. B., Bruce, C., Rozowsky, J. S., Zheng, D., Du, J., Korbel, J. O., … & Snyder, M. (2007). What is a gene, post-ENCODE? History and updated definition. *Genome Research*, 17(6), 669–681. doi:10.1101/gr.6339607

Goddard, M. E. (2009). Genomic selection: Prediction of accuracy and maximization of long-term response. *Genetics*, 136, 245–257. doi:10.1007/s10709-008-9308-0

Goddard, M. E., & Hayes, B. J. (2007). Genomic selection*. Journal of Animal Breeding and Genetics*, 124, 323-330. doi:10.1111/j.1439-0388.2007.00702.x

Grisart, B., Coppieters, W., Farnir, F., Karim, L., Ford, C., Berzi, P., … & Snell, R. (2002). Positional candidate cloning of a QTL in dairy cattle: identification of a missense mutation in the bovine DGAT1 gene with major effect on milk yield and composition. *Genome Research*, 12(2), 222–231. doi:10.1101/gr.224202

Hashim, H., & Al-Shuhaib, M. (2019). Exploring the potential and limitations of PCR-RFLP and PCR-SSCP for SNP detection*: A Review. Journal of Applied Biotechnology Reports*, 6(4), 137-144. doi:10.29252/JABR.06.04.02

Hayes, B. (2013). Overview of statistical methods for genome-wide association studies (GWAS). Genome-wide association studies and genomic prediction. *Methods in Molecular Biology*, 1019, 149–169. doi:10.1007/978-1-62703-447-0_6

Houba, P. H. J. & Te Pas, M. F. W. (2004). The muscle regulatory factors gene family in relation to meat production. In Te Pas, M.F.W.; Everts, M.E., Haagsman, H.P. (2004). Muscle development of livestock animals: physiology, genetics and meat quality. Oxfordshire: CABI, p.201-224.

Jena, K. K., & Mackill, D. J. (2008). Molecular markers and their use in marker-assisted selection in rice. *Crop Science*, 48(4), 1266. doi:10.2135/cropsci2008.02.0082

Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24(11), 1403–1405. doi:10.1093/bioinformatics/btn129

Keenan, K., McGinnity, P., Cross, T. F., Crozier, W. W., & Prodöhl, P. A. (2013). diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution*, 4(8), 782–788. doi:10.1111/2041-210x.12067

Kim, J. M., Choi, B. D., Kim, B. C., Park, S. S., & Hong, K. C. (2009). Associations of the variation in the porcine myogenin gene with muscle fibre characteristics, lean meat production and meat quality traits. *Journal of Animal Breeding and Genetics*, 126, 134-141. doi:10.1111/j.1439-0388.2008.00724.x.

Kurz, J. P., Yang, Z., Weiss, R. B., Wilson, D. J., Rood, K. A., Liu, G. E., & Wang, Z. (2018). A genome-wide association study for mastitis resistance in phenotypically well-characterized Holstein dairy cattle using a selective genotyping approach. *Immunogenetics,* 71(1), 35-47. doi:10.1007/s00251-018-1088-9

Liabin, L., Chen, X., & Chen, Z. (2019). Identification of key candidate genes in dairy cow in response to *Escherichia coli* mastitis by bioinformatical analysis. *Frontiers in genetics*, 10, 1251. doi:10.3389/fgene.2019.01251

Lozada-Soto, E. A., Maltecca, C., Lu, D., Miller, S., Cole, J. B., & Tiezzi, F. (2021). Trends in genetic diversity and the effect of inbreeding in American Angus cattle under genomic selection. *Genetics Selection Evolution*, 53, 50. doi:10.1186/s12711-021-00644-z

Marwal, A., & Gaur, R. K. (2020). Molecular markers: tool for genetic analysis. *Animal Biotechnology*, 353–372. doi:10.1016/b978-0-12-811710-1.00016-1

Meuwissen, T. H. E., Hayes, B. J., & Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, 157(4), 1819-1829. doi:10.1093/genetics/157.4.1819

Meuwissen, T. H. E., Hayes, B. J., & Goddard M. E. (2016). Genomic selection: A paradigm shift in animal breeding. *Animal Frontiers*, 6, 6-14. doi:10.2527/af.2016-0002

Moré, D. D., Cardoso, F. F., Mudadu, M. A., Malagó-Jr, W., Gulias-Gomes, C. C., Sollero, B. P., … & Regitano, L. C. A. (2019). Network analysis uncovers putative genes affecting resistance to tick infestation in Braford cattle skin. *BMC Genomics*, 20, 998. doi:10.1186/s12864-019-6360-3

Moretti, R., Soglia, D., Chessa, S., Sartore, S., Finocchiaro, R., Rasero, R., & Sacchi, P. (2021). Identification of SNPs associated with somatic cell Score in candidate genes in Italian Holstein Friesian bulls. *Animals*, 11(2), 366. doi:10.3390/ani11020366

Mortezaei, Z., & Tavallaei, M. (2021). Novel directions in data pre-processing and genome-wide association study (GWAS) methodologies to overcome ongoing challenges. *Informatics in Medicine Unlocked*, 24, 100586. doi:10.1016/j.imu.2021.100586

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

Mota, R. R., Guimarães, S. E. F., Fortes, M. R. S., Hayes, B., Silva, F. F., Verardo, L. L., … & Moore, S. (2017a). Genome-wide association study and annotating candidate gene networks affecting age at first calving in Nellore cattle. *Journal of Animal Breeding and Genetics*, 134(6), 484–492. doi:10.1111/jbg.12299

Mota, R. R., Silva, F. F., Lopes, P. S., Tempelman, R. J., Sollero, B. P., Aguilar, I., & Cardoso, F. F. (2017b). Analyses of reaction norms reveal new chromosome regions associated with tick resistance in cattle. *Animal*, 12(02), 205–214. doi:10.1017/s1751731117001562

Moyen, C., Goudenege, S., Poussard, S., Sassi, A. H., Brustis, J.-J., & Cottin, P. (2004). Involvement of micro-calpain (CAPN 1) in muscle cell differentiation. *The International Journal of Biochemistry & Cell Biology*, 36(4), 728–743. doi:10.1016/s1357-2725(03)00265-6

Nadeem, M. A., Nawaz, M. A., Shahid, M. Q., Dogan, Y., Comertpay, G., Yildiz, M., … & Baloch, F. S. (2017). DNA molecular markers in plant breeding: current status and recent advancements in genomic selection and genome editing. *Biotechnology & Biotechnological Equipment*, 32(2), 261–285. doi:10.1080/13102818.2017.1400401

Ortega, M. S., Denicol, A. C., Cole, J. B., Null, D. J., & Hansen, P. J. (2016). 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(3), 288–297. doi:10.1111/age.12420

Page, B. T., Casas, E., Quaas, R. L., Thallman, R. M., Wheeler, T. L., Shackelford, S. D., … & Smith, T. P. (2004). Association of markers in the bovine CAPN1 gene with meat tenderness in large crossbred populations that sample influential industry sires. *Journal of Animal Science*, 82(12): 3474-3481. doi: 10.2527/2004.82123474x

Pimentel, E. C. G., Bauersachs, S., Tietze, M., Simianer, H., Tetens, J., Thaller, G., … & König, S. (2010). Exploration of relationships between production and fertility traits in dairy cattle via association studies of SNPs within candidate genes derived by expression profiling. *Animal Genetics*, 42(3), 251–262. doi:10.1111/j.1365-2052.2010.02148.x

Platten, J. D., Cobb, J. N., & Zantua, R. E. (2019). Criteria for evaluating molecular markers: Comprehensive quality metrics to improve marker-assisted selection. *PLOS ONE*, 14(1), e0210529. doi:10.1371/journal.pone.0210529

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/

Rubtsov, A. M., & Lopina, O. D. (2000). Ankyrins. *FEBS Letters*, 482, 1-5.

Singh, U., Deb, R., Alyethodi, R. R., Alex, R., Kumar, S., Chakraborty, S., … & Sharma, A. (2014). Molecular markers and their applications in cattle genetic research: A review. *Biomarkers and Genomic Medicine*, 6(2), 49-58. doi:10.1016/j.bgm.2014.03.001

Song, M., He, Y., Zhou, H., Zhang, Y., Li, X., & Yu, Y. (2016). Combined analysis of DNA methylome and transcriptome reveal novel candidate genes with

susceptibility to bovine *Staphylococcus aureus* subclinical mastitis. *Scientific Reports*, 6, 29390. doi:10.1038/srep29390

Soria L. A., Corva, P. M., Huguet, M. J., Miño, S., & Miquel, M. C. (2010). Bovine calpain (CAPN1) gene polymorphisms in Brangus and Brahman bulls. *Journal of Basic and Applied Genetics*, 21(1), 61-69.

Te Pas, M. F. W. & Soumillion, A. (2001). Improvement of livestock breeding strategies using physiologic and functional genomic information of the muscle regulatory factors gene family for skeletal muscle development. *Current Genomics*, 2, 285-304.

Te Pas, M. F. W., Soumillion, A., Harders, F. L., Verburg, F. J., Van den Bosch, T. J., Galesloot, P., & Meuwissen, T.H. (1999). Influences of myogenin genotypes on birth weight, growth rate, carcass weight, backfat thickness, and lean weight of pigs. *Journal of Animal Science*, 77, 2352-2356. doi:10.2527/1999.7792352x.

Tian, D., Wang, P., Tang, B., Teng, X., Li, C., Liu, X., … & Zhang, Z. (2019). GWAS Atlas: a curated resource of genome-wide variant-trait associations in plants and animals. *Nucleic Acids Research*. doi:10.1093/nar/gkz828

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, 307. doi:10.1007/s11250-021-02757-5

Uitterlinden, A. (2016). An Introduction to Genome-Wide Association Studies: GWAS for Dummies. *Seminars in Reproductive Medicine*, 34(04), 196–204. doi:10.1055/s-0036-1585406

Ujan, J.A., Zan, L.S., Shengjuan, W., Adoligbe, C. & Wang, H.B. (2011). Meat tenderness and water holding capacity are associated with a 959 A G mutation in the MYOG gene of Chinese indigenous cattle. *African Journal of Biotechnology*, 10, 5654-5660.

White, S. N., Casas, E., Wheeler, T. L., Shakelfold, S. D., Koohmaraie, M., Riley, D.G., … & Smith, T. P. L. (2005). A new single nucleotide polymorphism in CAPN1 extends the current tenderness marker test to include cattle of Bos indicus, Bos taurus, and crossbred descent. *Journal of Animal Science*, 83(9): 2001-2008. doi: 10.2527/2005.8392001x

Xia, X., Qu, K., Li, F., Jia, P., Chen, Q., Chen, N., … & Lei, C. (2019). Abundant Genetic Diversity of Yunling Cattle Based on Mitochondrial Genome. *Animals*, 9(9), 641. doi:10.3390/ani9090641

Yang, W., Kang, X., Yang, Q., Lin, Y., & Fang, M. (2013). Review on the development of genotyping methods for assessing farm animal diversity. *Journal of Animal Science and Biotechnology*, 4(1), 2. doi:10.1186/2049-1891-4-2

Yin, T., & König, S. (2016). Genomics for phenotype prediction and management purposes. *Animal Frontiers*, 6(1), 65–72. doi:10.2527/af.2016-0010

This chapter was published in *Tropical Animal Health and Production*

CAPÍTULO 3. ALLELIC AND GENOTYPIC FREQUENCIES FOR LOCI ASSOCIATED WITH MEAT QUALITY IN MEXICAN BRAUNVIEH CATTLE

Tropical Animal Health and Production (2021) 53:307
https://doi.org/10.1007/s11250-021-02757-5

REGULAR ARTICLES

Allelic and genotypic frequencies for loci associated with meat quality in Mexican Braunvieh cattle

Mitzilin Zuleica Trujano-Chavez[1] · Jonathan E. Valerio-Hernández[1] · Rufino López-Ordaz[1] · Paulino Pérez-Rodríguez[2] · Agustín Ruíz-Flores[1]

Received: 23 December 2020 / Accepted: 2 May 2021

Abstract
The objective was to estimate allelic and genotypic frequencies for loci associated with meat quality in a Mexican population of Braunvieh cattle. Information was obtained from 300 animals genotyped with the Genomic Profile Bovine LD chip of 30K and 50K SNPs. After the final edition, including quality control, the data contained information for 12 loci of the CAPN1, CAPN3, CAPN5, CAPN14, DGAT1, DGAT2, TG, ANK1, and MADH3 genes. Allelic and genotypic frequencies and Hardy-Weinberg equilibrium were estimated with the Cervus 3.0.7 software. The studied population markers were in Hardy-Weinberg equilibrium, except for those associated with CAPN1, DGAT1, and MADH3. Frequencies higher than those reported for other breeds were found for genotypes associated with meat softness, higher marbling score, lower quantity of saturated fatty acids, and lower shear force (CAPN1 and DGAT2). There were similarities with frequencies reported for *Bos taurus* breeds for the CAPN3 and TG genes. For the DGAT1 and ANK1 genes, the frequencies of the desired genotypes were low. A marker for DGAT1 and another for MADH3 were monomorphic. The results of this study are encouraging in terms of the potential of the Braunvieh population studied for breeding programs aiming to increase meat quality. The breed has strengths that could be used either by crossbreeding to generate heterozygous animals or by selection to increase frequencies of valuable alleles.

Keywords Meat quality · DGAT1 · TG · CAPN1 · Genomic characterization · Hardy-Weinberg equilibrium

Introduction

Good quality beef is highly demanded. Meat quality is determined partially by visual, sensorial, and nutritional characteristics along with food safety and animal welfare that make it attractive for the consumer (Oliveira et al. 2014). Some of these characteristics are high marbling score, softness, juiciness, and tenderness (Ruiz et al. 2003).

Marker-assisted genomic selection (GWAS) could be a suitable option for achieving genetic progress in meat quality. This trait, as well as others of economic interest, according to Grisart et al. (2002), is of polygenic type and is under the control of several loci for quantitative traits (QTL). A long generational interval limits traditional genetic improvement in beef cattle. Moreover, meat quality traits are usually measured in slaughtered animals, which delay the genetic progress and make the selection of young animals difficult.

If a gene is significantly associated with a trait or within a region harboring a QTL, it is called candidate gene (Long 2001). Significant effects for the CAPN1 gene for meat softness and marbling were reported in Brangus (Mazzucco et al. 2010), Canchim cattle (Curi et al. 2010), and Nelore (Tizioto et al. 2014). Additionally, Oliveira et al. (2014) in Nelore cattle found a close association of ANK1 with meat color and DGAT1 with increased marbling. The close association of marbling with DGAT1 was also reported in Nelore and Angeln cattle (Grisart et al. 2002; Sanders et al. 2006; Souza et al. 2010). In another study, Rubtsov and Lopina (2000) documented an important effect of ANK1 polymorphism on meat softness.

The frequency of polymorphisms associated with meat quality has been characterized genomically (allelic and genotypic frequencies for candidate genes) in Nelore populations in Brazil (Oliveira et al. 2014), Brangus in Argentina

✉ Agustín Ruíz-Flores
aruizf@chapingo.mx

[1] Universidad Autónoma Chapingo, 56227 Chapingo, State of Mexico, Mexico

[2] Colegio de Postgraduados, 56264 Montecillos, State of Mexico, Mexico

Published online: 06 May 2021

(Mazzucco et al. 2010), and Braunvieh in Brazil (Curi et al. 2010). However, to our knowledge, this kind of study has not been conducted with Braunvieh cattle in Mexico. The objective of this study was to estimate allelic and genotypic frequencies for loci of candidate genes associated with meat quality in a Mexican Braunvieh population.

Materials and methods

Genotypic data

Hair samples were collected from 300 Braunvieh animals (236 female, 64 male) born from 2001 to 2016 on purebred farms belonging to the Mexican Association of Braunvieh Breeders from five herds in Eastern, Central, and Western Mexico. The samples were genotyped at GeneSeek (Lincoln, NE, USA, http://genomics.neogene.com). The chip used was Genomic Profile Bovine LD with 30,000 and 50,000 SNP markers for 150 different animals each.

Loci identification

Markers or loci were identified using the genomic information available for *Bos taurus* at Gene (2020). The loci-gene allocation criterion was that the marker position was within the gene's allocation position range for the same chromosome. Only loci available in the genotyped samples were included.

Twelve loci of nine candidate genes reported in the literature to be associated with meat quality were used; these genes are shown in Table 1. Information of 30K and 50K genotypes was unified for CAPN1, CAPN5, CAPN14, DGAT2, TG, ANK1, and MADH3, each with one assigned marker, DGAT1 with two markers, and CAPN3 with three. The loci-gene allocation criterion was that the marker position was within the same chromosome's gene position range.

Data analysis

For CAPN1 and DGAT1 analysis, information from 300 animals was used, while for the other seven genes, the analysis included 150 animals because the markers of the 30K marker chip did not contain the other loci in the study. Allelic and genotypic frequencies, as well as the Chi-squared test for Hardy-Weinberg equilibrium, were obtained with the Cervus 3.0.7 software package (Kalinowski et al. 2007).

The equation (Falconer and Mackay 1996) used by the software to test Hardy-Weinberg equilibrium is:

$$p^2 + 2pq + q^2 = p + q = 1$$

where p^2 is AA homozygous genotypic frequency, $2pq$ is heterozygote genotypic frequency (Aa), and q^2 is aa homozygous genotypic frequency. Also, p and q represent the A and a allelic frequencies.

Results and discussion

The results of the characterization for meat quality candidate genes in the Braunvieh population studied are shown in Table 2. The allelic (alleles A and a) and genotypic (AA, Aa, and aa), general information of each marker, and its corresponding Chi-squared test to determine whether the population was or not in Hardy-Weinberg equilibrium (HW) are presented. The level of significance used in the test was 0.05; thus, values above 0.05 indicate that the population is in Hardy-Weinberg equilibrium.

Eight of the markers were in HW equilibrium (Table 2). However, markers associated with genes of major importance, such as CAPN1 and DGAT1, were not in HW equilibrium, and it could be due to either selection of the animals or sample size.

The studied population had a larger frequency of the desired genotype in the CAPN1 gene (AA frequency = 0.25) than other beef cattle breeds. In the literature it has been documented that genotypic variants (AA, Aa, and aa) did not significantly affect meat color. However, it has been reported that genotype AA is associated with higher marbling scores than the other genotypes (Li et al. 2013). These authors found frequencies of 0.00 to 0.11 for this genotype in *B. taurus* breeds (Angus had the highest frequency). The frequency for this genotype in Braunvieh is 0.25, Table 2, higher than in any previously mentioned breeds. According to Bonilla et al. (2010), in *Bos indicus*× *Bos taurus* cattle, the situation is similar: the frequency of genotype AA is on average 0.15. Similarly, genotype AA increases meat softness (Mazzucco et al. 2010). Therefore, allele A frequency in cattle populations should be as high as possible. In the present study, the results (0.5) were similar to those obtained (0.49) in Brangus (Mazzucco et al. 2010). In Brazilian Nelore and Braunvieh (Curi et al. 2010), genotype AA frequencies were essentially zero (0.000081 and 0.04489), results different from those of our study. These authors hypothesized that their values were due to the relatively small sample size used, 114 Nelore and 15 Braunvieh.

The Braunvieh population studied had similarities with other *B. taurus* breeds for the desirable genotypic frequencies of the gene CAPN3 (frequency range between 0.27 and 0.71). According to Barendse et al. (2008), the valuable genotype of the CAPN3 gene for greater softness is AA. These authors characterized the Red Belmont and Santa Gertrudis (frequencies for genotype AA, 0.44 and 0.32). Our study's frequency for this genotype was 0.27, close to that reported for Red Belmont and Santa Gertrudis. However, Brahman has a lower

Table 1 Frequencies of homozygous and heterozygous genotypes reported in studies on beef cattle populations for genes associated with meat quality

Gene	Symbol	Chrom	Genetic Group	F_{Ht}	F_{Ho}	Reference
Calpain 1 [1, 2, 3, 5, 14]	CAPN1	29	Brangus	0.45	0.55	Mazzucco et al. 2010
			B. taurus×B. indicus	0.40	0.60	Bonilla et al. 2010
			Angus	0.49	0.51	Li et al. 2013
			Charolais	0.21	0.79	
			Hereford	0.06	0.94	
			Simmental	0.33	0.67	
			Nelore	0.02	0.98	Curi et al. 2010
			Braunvieh	0.13	0.87	
			Canchim	0.17	0.83	
			Angus×Hereford	0.41	0.59	Corva et al. 2007
Calpain 3 [2]	CAPN3	10	Brahman	0.42	0.58	Barendse et al. 2008
			Red Belmont	0.21	0.79	
			Santa Gertrudis	0.31	0.69	
Calpain 5 [4, 2]	CAPN5	15	Nelore	N/R	N/R	Tizioto et al. 2013
			Blonde d'Aquitaine	N/R	N/R	De Oliveira et al. 2017
			Charolais	N/R	N/R	Ramayo-Caldas et al. 2016
			Limousin	N/R	N/R	
Calpain 14 [4]	CAPN14	11	Wagyu×Limousin	N/R	N/R	Jiang et al. 2009
Thyroglobulin [5, 7, 8]	TG	14	Simmental	0.48	0.52	Gan et al. 2008
			Angus	0.24	0.76	
			Hereford	0.29	0.71	
			Charolais	0.07	0.93	
			Limousin	0.42	0.58	
			Jinnan	0.43	0.57	
			Nelore	0.00	1.00	Fortes et al. 2009
			Rubia Gallega×Nelore	0.04	0.96	
			Brangus	0.40	0.60	
			Braunvieh	0.39	0.61	
Diacylglycerol o-acyltransferase 1 [5, 3, 6]	DGAT1	14	Angus	0.98	0.02	Li et al. 2013
			Charolais	1.00	0.00	
			Limousin	1.00	0.00	
			Simmental	1.00	0.00	
			Luxi	0.22	0.78	Yuan et al. 2013
			Qinchuan	0.24	0.76	
			Belgian blue	0.07	0.93	Pannier et al. 2010
			Blonde d'Aquitaine	0.33	0.63	
			Salers	0.00	1.00	
			Friesian	0.23	0.77	
Dyacilglycerol o-acyltransferase 2 [13, 5, 2]	DGAT2	15	Angus	0.00	1.00	Li et al. 2009
			Jersey	N/R	N/R	Dunner et al. 2013
			Holstein	N/R	N/R	
			Simmental	N/R	N/R	
			Avileña-Negra Ibérica	N/R	N/R	
			Piedmontese	N/R	N/R	
			Marchigiana	N/R	N/R	
			Limousin	N/R	N/R	
			Charolais	N/R	N/R	
Malate dehydrogenase 3 [2]	MADH3	10	*Bos Taurus*	N/R	N/R	Dunner et al. 2013
Ankyrin 1 [1, 5, 2, 9, 10, 11, 12, 14]	ANK 1	27	Nelore	0.47	0.53	Oliveira et al. 2014
			Charolais	0.13	0.87	Horodyska et al. 2015

Table 1 (continued)

Gene	Symbol	Chrom	Genetic Group	F_{Ht}	F_{Ho}	Reference
			Angus	0.19	0.81	Aslan et al. 2010
			Charolais	0.29	0.71	
			Limousin	0.36	0.64	

Symbol, Symbol of the gene; *Chrom*, Chromosome; F_{Ht}, Heterozygous genotype frequency; F_{Ho}, homozygous genotype frequency; *N/R* = Not reported. Numeric superscripts in the names of genes indicate the associated meat quality characteristics: [1] Meat color, [2] Tenderness, [3] Shear force, [4] Amount of intramuscular fat, [5] Marbling, [6] Fat color, [7] Loin-eye area, [8] Rib-eye area, [9] Meat texture, [10] Juiciness, [11] Sarcomere length, [12] Cooking losses, [13] Fatty acid content, and [14] Softness

frequency, 0.19, than Braunvieh for this genotype (Barendse et al. 2008).

In the literature consulted, there were not found reports for allelic and genotypic frequencies for CAPN5, CAPN14, and MADH3 genes in beef cattle. However, in literature, there is ample evidence of association with meat quality traits, Table 1. In the present study, variability was found (Table 2) for CAPN5 genotypes. Hopefully, future research would allow a better characterization of Braunvieh regarding its allelic and genotypic frequencies. CAPN14 in Wagyu×Limousin crossbred cattle was found associated with seam fat (Jiang et al. 2009), as in Nelore (Tizioto et al. 2013). In the present study, for Braunvieh was nearly homozygous (aa = 0.93), the genetic variation is small. This could be either positive or negative, depending on further studies that elucidate the best genotype for CAPN14. Dunner et al. (2013) found an association of MADH3 with postmortem myofibrillar fragmentation (tenderness). Frequencies for this gene were in a situation like that of the CAPN14 gene (AA = 1). Findings of future research will determine the convenience of this homozygosity.

Genotypic frequencies for the DGAT1 gene reported in literature have been different among breeds. Most of them had high proportions of the favorable genotype for marbling. According to our results, it was not present in Braunvieh. Li et al. (2013) reported an association of the DGAT1 gene with marbling in four European breeds. The heterozygous Aa genotype was associated with the best phenotype. In the same study, these authors reported for all four breeds, heterozygous frequencies above 0.98; therefore, *B. taurus* breeds would have the best DGAT1 genotype for marbling. This is contrary to our study's estimated frequencies, which coincide with those reported by Yuan et al. (2013) in Chinese commercial cattle with Aa frequencies below 0.24. Pannier et al. (2010) in Belgian Blue and Fortes et al. (2009) in Nelore, reported low

Table 2 Allelic (A and a) and genotypic (AA, Aa, and aa) frequencies for loci associated with meat quality characteristics, and Hardy-Weinberg Equilibrium (HW) in the Braunvieh population studied

SG	Chrom	SNP Marker	Position	Allele		Genotype			HW
				A	a	AA	Aa	aa	
CAPN1	29	UA-IFASA-1370	44085769	0.50*	0.50	0.25*	0.50	0.25	b
CAPN3	10	BovineHD1000011720	37860216	0.71*	0.29	0.51*	0.41	0.08	a
		BovineHD1000011723	37870339	0.47*	0.53	0.22*	0.50	0.28	a
		Hapmap47063-BTA-62293	37852123	0.29*	0.71	0.09*	0.41	0.50	a
CAPN5	15	BovineHD1500016485	57266887	0.40	0.60	0.16	0.48	0.36	a
CAPN14	11	ARS-BFGL-NGS-10829	68687376	0.03	0.97	0.00	0.07	0.93	a
DGAT1	14	ARS-BFGL-NGS-4939	1801116	0.94	0.06	0.88	0.11*	0.00	b
		Hapmap31987-BTC-062044	8425401	1.00	0.00	1.00	0.00*	0.00	b
DGAT2	15	UA-IFASA-5331	55185032	0.64	0.36	0.41*	0.46*	0.13	a
TG	14	BovineHD1400002314	8406694	0.64	0.36	0.41	0.46*	0.13	a
ANK1	27	ARS-BFGL-NGS-88722	36342691	0.23*	0.77	0.05*	0.36	0.59	a
MADH3	10	BTB-00410322	13990943	1.00	0.00	1.00	0.00	0.00	b

SG, Symbol of the gene; *Chrom*, Chromosome. The frequencies with the superscript (*) indicates the desired alleles and/or genotypes by gene (improvement in meat quality). In the HW column, superscript [a] indicates there were non-significant differences ($P > 0.05$) between the expected and the observed values in the Chi-squared test; this means the population is in Hardy-Weinberg equilibrium for the corresponding marker. On the contrary, [b] indicates the population is not in HW equilibrium

frequencies for the heterozygous Aa genotype, results like ours (0.07 in Belgian Blue and Nelore *vs* 0.00-0.11 in Braunvieh). Genetic improvement in Belgian Blue oriented to lean meat production could limit genotypes that code for higher production and deposit of fat, possibly explaining these DGAT1 frequencies in this breed.

Though less studied, the DGAT2 gene has significant effects on quality characteristics related to fat in *B. taurus* breeds. The Braunvieh population studied has the genetic potential for less seam fat. According to Li et al. (2009), the desired genotypes, in this case, were AA and Aa, associated with a low amount of seam fat. These authors studied three native Chinese breeds (Luxi, Jinnan, and Qinchuan), three crossbred (Charolais×Fuzhou, Limousin×Luxi, and Simmental×Jinan) and one *B. taurus* pure breed population (Angus). They found that genotype aa had frequencies above 0.98, results contrary to the 0.13 of our study, which means the total frequency of desirable genotypes is 0.87.

Braunvieh cattle exhibited frequencies for the desired TG genotype, Aa, like other *B. taurus* breeds. The TG gene is one of the most studied in beef cattle. According to Oliveira et al. (2014), in a study carried out in Brazil, the Braunvieh population had frequencies for Aa like those obtained in our study (0.39 vs 0.46). Gan et al. (2008) studied eight breeds; they reported that the desired genotype was the heterozygous, Aa, including the effects of six traits: live weight, carcass weight, carcass yield, seam fat, marbling, and loin- and rib-eye area. In the same study, the breed characterization showed that Simmental, Jinnan, and Limousin had the best frequencies (above 0.42), while Charolais had only 0.07 for the heterozygous genotypes. Our results in Braunvieh were like those reported for the first three breeds but higher than for Jinnan and Limousin (0.46 *vs* 0.43 and 0.42). On the other hand, Zhang et al. (2015), who studied the same breeds as Gan et al. (2008), found different results for TG, with high frequencies for homozygous genotypes, possibly due to the small number of animals used for estimations, on average 34 of each breed.

The favorable genotype for ANK1 (AA) is rare in Braunvieh compared with other breeds. The ANK1 gene, like TG, is one of the most studied and, according to Oliveira et al. (2014), the most valuable allele, considering intramuscular fat, rib-eye area, and cooking loss, is A. This means the best genotype is AA. The same authors reported a frequency of 0.2 for AA in Nelore, similar to findings in Charolais with a frequency of 0.32 (Horodyska et al. 2015). However, the differences in frequency for this gene in other breeds with Braunvieh are notorious, 0.2 vs 0.05.

In general, the Braunvieh population studied had higher frequencies of the desired genotypes than those of other breeds for CAPN1 and DGAT2 genes, mainly associated with fat content in meat. It also presents similarities to CAPN3 and TG gene frequencies for most of *B. taurus* breeds, while for the DGAT1 and ANK1 genes, the frequencies of the most valuable genotypes were lower than those of other breeds. For the genes for which information was not found in literature, comparisons could not be made: CAPN5 presented a variation in the genotypes. In contrast, CAPN14 and MDH3 were monomorphic and homozygous; their interpretation is pending future research for valid conclusions.

Genetic frequencies indicate the genetic potential of a breed for economically important traits associated with meat quality. Only two of the 12 markers were monomorphic, indicating genetic variation in the studied population. The frequencies were high for valuable genotypes of genes associated with meat quality traits such as greater softness, marbling, lower amount of saturated fatty acids and shear force (CAPN1 and DGAT2). TG had similarities with other breeds, with high frequencies. This means that the population studied is potentially superior in the loin- and rib-eye area. The results were not favorable for genes associated with the yellow color of fat and less marbling; the desired genotype had frequencies below 0.07 (DGAT1 and ANK1).

The results in this study for meat quality traits, widely recognized as marbling and softness, were encouraging for the potential use of Braunvieh in beef cattle breeding programs. Selection could be used to increase alleles' frequency with high individual value to generate the homozygous genotypes favorable for meat quality traits.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11250-021-02757-5.

Acknowledgements The authors thank the Asociación Mexicana de Criadores de Ganado Suizo de Registro for allowing the use of their databases and the cooperation of the breeders for this study. Thanks also are extended to the Consejo Nacional de Ciencia y Tecnología, Mexico, for the financial support of the first author during her Master of Science studies.

Author contribution MZTC wrote the first version of the manuscript, JEVH and PPR assisted MZTC during the analysis, RLO and ARF conceived the project, all the authors contributed to obtain the final version, ARF served as the correspondence author.

Funding This work was supported by a grant from the Universidad Autónoma Chapingo, Mexico (DGIP-166701012).

Declarations

Conflict of interest We declare no conflict of interest with any financial organization regarding the material discussed in the manuscript.

Statement of animal rights All applicable international, national, and/or institutional guidelines for the care and use of animals were followed. The manuscript does not contain clinical studies or patient data.

References

Aslan, O., Sweeny, T., Mullen, A.M. and Hamill, R.M., 2010. Regulatory polymorphisms in the bovine Ankyrin 1 gene promoter are associated with tenderness and intramuscular fat content, BMC Genetics, 11, 111-1124.

Barendse, W., Harrison, B.E., Bunch, R.J. and Thomas, M.B., 2008. Variation at the Calpain 3 gene is associated with meat tenderness in zebu and composite breeds of cattle, BMC Genetics, 9(41), 1-8.

Bonilla, C.A., Rubio, M.S., Sifuentes, A.M., Parra-Bracamonte, G.M., Arellano, V.W., Méndez, M.R.D., Berruecos, J.M. and Ortiz, R., 2010. Association of CAPN1 316, CAPN1 4751 and TG5 markers with bovine meat quality traits in Mexico, Genetics and Molecular Research 9(4), 2395-2405.

Corva, P., Soria, L., Schor, A., Villarreal, E., Pérez, C.M., Motter, M., Mezzadra, C., Melucci, L., Miquel, C., Paván, E., Depetris, G., Santini, F. and Naón, J.G., 2007. Association of CAPN1 and CAST gene polymorphisms with meat tenderness in *Bos taurus* beef cattle from Argentina, Genetics and Molecular Biology, 30(4), 1064-1069.

Curi, R.A., Chardulo, L.A.L., Giusti, J., Silveira, A.C., Martins, C.L. and de Oliveira, H.N., 2010. Assessment of GH1, CAPN1, and CAST polymorphisms as markers of carcass and meat traits in Bos indicus and Bos taurus-Bos indicus cross beef cattle, Meat Science, 86, 915-920.

De Oliveira, S.R.M., Bonvino, S.N., de Oliveira, B.F., Ferreira, M.G.C., Matos T.C., Goncalves, J.N.S.C., Baldi, F., Boligon, A.A., Zerlotti Mercadante, M.E., Lourenco, D.L., Misztal, I. and de Albuquerque, L.G., 2017. Genome-wide association study for carcass traits in an experimental Nelore cattle population, Plos One, 12(1), e0169860.

Dunner, S., Sevane, N., García, D., Cortés, O., Valentini, A., Williams, J.L., Mangin, B., Cañón, J., Levéziel, H. and the GeMQual Consortium, 2013. Association of genes involved in carcass and meat quality traits in 15 European bovine breeds, Livestock Science, 154, 34-44.

Falconer, D.S. and Mackay, T.F.C., 1996. Introduction to Quantitative Genetics. (Pearson, Prentice Hall, New Jersey).

Fortes, M.R.S., Curi, R.A., Chardulo, L.A.L., Silveira, A.C., Assumpção, M.E.O.D., Visintin, J.A. and de Oliveira, H.N., 2009. Bovine gene polymorphisms related to fat deposition and meat tenderness, Genetics and Molecular Biology, 32(1), 75-82.

Gan, Q.F., Zhang, L.P., Li, J.Y., Hou, G.Y., Li, H.D., Gao, X., Ren, H.Y., Chen, J.B. and, Xu, S.Z., 2008. Association analysis of thyroglobulin gene variants with carcass and meat quality traits in beef cattle, Journal of Applied Genetics, 49, 251-255.

Gene., 2020. National Library of Medicine (US). National Center for Biotechnology Information (NCBI). https://www.ncbi.nlm.nih.gov/gene. Accessed 19 March, 2020.

Grisart, B., Coppieters, W., Farnir, F., Karim, L., Ford, C., Berzi, P., Cambisano N., Mni, M., Reid, S., Simon, P., Spelman, R., Georges, M. and Snell, R., 2002. Positional candidate cloning of a QTL in dairy cattle: identification of a missense mutation in the bovine DGAT1 gene with major effect on milk yield and composition, Genome Research, 12, 222–231.

Horodyska, J., Sweeney, T., Ryan, M. and Hamill, R.M., 2015. Novel SNPs in the Ankiryn 1 gene and their association with beef quality traits, Meat Science, 108, 88-96.

Jiang, Z., Michal, J.J., Chen, J., Daniels, T.F., Kunej, T., Garcia, M.D., Gaskins, C.T., Busboom, J.R., Alexander, L.J., Wright Jr., R.W. and MacNeil, M.D., 2009. Discovery of novel genetic networks associated with 19 economically important traits in beef cattle, International Journal of Biological Sciences, 5(6), 528-542.

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

Li, J., Xu, X., Zhang, Q., Wang, X., Deng, G., Fang, X., Gao, X., Ren, H. and Xu, S., 2009. Association between single nucleotide polymorphisms in the dgat2 gene and beef carcass and quality traits in commercial feedlot steers, Asian-Australasian Journal of Animal Sciences 7, 943-954.

Li, X., Ekerljung, M., Lundström, K. and Lundén, A., 2013. Association of polymorphisms at DGAT1, leptin, SCD1, CAPN1 and CAST genes with color, marbling and water holding capacity in meat from beef cattle populations in Sweden, Meat Science, 94, 153-158.

Long, A., 2001. Candidate Gene, In: Encyclopedia of Genetics, S. Brenner, and J. H. Miller, Editor(s). (Academic Press, Cambridge, MA).

Mazzucco, J.P., Melucci, L.M., Villareal, E.L., Mezzadra, C., Soria, L., Corva, P., Motter, M.M., Schor, A. and Miquel, M.C., 2010. Effect of ageing and μ-calpain markers on meat quality from Brangus steers finished on pasture, Meat Science, 86(3), 878-882.

Oliveira, B.B., Abdallah, C.R., Baldi, F., Braga, F.F.L., Fernandes, A.W.B., Galvão, A.L., Nunes, O.H. and Loyola, C.L.A., 2014. Polymorphisms in candidate genes and their association with carcass traits and meat quality in Nelore cattle, Pesquisa Agropecuária Brasileira, 49(5), 364-371.

Pannier, L., Mullen, A.M., Hamill, R.M., Stapleton, P.C. and Sweeney T., 2010. Association analysis of single nucleotide polymorphisms in DGAT1, TG and FABP4 genes and intramuscular fat in crossbred *Bos taurus* cattle, Meat Science, 85, 515-518.

Ramayo-Caldas, Y., Renand, G., Ballester, M., Saintilan, R. and Rocha D., 2016. Multi-breed and multi-trait co-association analysis of meat tenderness and other meat quality traits in three French beef cattle breeds, Genetics Selection Evolution, 48(37), 1-9.

Rubtsov, A.M. and Lopina, O.D., 2000. Ankyrins, FEBS Letters 482, 1-5.

Ruiz de H, F., Miguel, E., Onega, E. and Blázquez B., 2003. Changes in meat quality characteristics of bovine meat during the first 6 days *postmortem*, Meat Science, 65, 1439-1446.

Sanders, K., Bennewitz, J., Reinsch, N., Thaller, G., Prinzenberg, E.-M., Kühn, C. and Kalm, E., 2006. Characterization of the DGAT1 mutations and the CSN1S1 promoter in the German Angeln dairy cattle population, Journal of Dairy Science, 89, 3164-3174.

Souza, F.R.P., Mercadante, M.E.Z., Fonseca, L.F.S., Ferreira, M.L.S., Regatieri, I.C., Ayres, D.R., Tonhati, H., Silva, S.L., Razook, A.G. and Albuquerque, L.G., 2010. Assessment of DGAT1 and LEP gene polymorphisms in three Nelore (*Bos indicus*) lines selected for growth and their relationship with growth and carcass traits, Journal of Animal Science, 88, 435-441.

Tizioto, P.C., Decker, J.E., Taylor, J.F., Schnabel, R.D., Mudadu, M.A., Silva, F.L., Mourão, G.B., Coutinho, L.L., Tholon, P., Sonstegard, T.S., Rosa, A.N., Alencar, M.M., Tullio, R.R., Medeiros, S.R., Nassu, R.T., Feijó, G.L.D., Silva, L.O.C., Torres, R.A., Siqueira, F., Higa, R.H. and Regitano, L.C.A., 2013. Genome scan for meat quality traits in Nelore beef cattle, Physiological Genomics 45, 1012-1020.

Tizioto, P.C., Gromboni, C.F., de Araujo, N.A.R., de Souza, M.M., Mudadu, M.A., Tholon, P., Rosa, A.N., Tullio, R.R., Medeiros, S.R., Nassu, R.T., and Regitano, L.C.A., 2014. Calcium and potassium content in beef: Influences on tenderness and associations with molecular markers in Nelore cattle, Meat Science, 96(1), 436-440.

Yuan, Z., Li, J., Li, J., Gao, X., Gao, H. and Xu, S., 2013. Effects of DGAT1 gene on meat and carcass fatness quality in Chinese commercial cattle, Molecular Biology Reports, 40, 1947-1954.

Zhang, L.P., Gan, Q.F., Hou, G.Y., Gao, H.J., Li, J.Y. and Xu, S.Z., 2015. Investigation of TG gene variants and their effects on growth, carcass composition, and meat quality traits in Chinese steers, Genetics and Molecular Research, 14(2), 5320-5326.

This chapter was sent to *Reproduction in Domestic Animals*

CAPÍTULO 4. GENETIC DIVERSITY FOR REPRODUCTIVE TRAITS IN BRAUNVIEH CATTLE DETERMINED WITH SNP MARKERS

Reproduction in Domestic Animals

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

Journal:	*Reproduction in Domestic Animals*
Manuscript ID	RDA-OR-Oct-2021-0471
Wiley - Manuscript type:	Original Article
Date Submitted by the Author:	29-Oct-2021
Complete List of Authors:	Trujano-Chavez, Mitzilin Zuleica; Universidad Autonoma Chapingo Ruiz-Flores, Agustin; Universidad Autonoma Chapingo López-Ordaz, Rufino; Universidad Autonoma Chapingo Pérez-Rodríguez, Paulino; Colegio de Postgraduados
Subject Area:	Animal breeding < General reproduction, Genetics < General reproduction, cattle < Species:
Editor Preference Recommended:	Prof. Dr. E. Martinez-Garcia, Murcia, Spain
Editor Preference Opposed:	Prof. Dr. F. Almeida, Belo Horizonte, Brazil

RDA Manuscript Proof

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

Mitzilin Zuleica Trujano-Chavez[a], Agustín Ruíz-Flores[a], Rufino López-Ordaz[a], Paulino Pérez-Rodríguez[b]

[a] *Posgrado en Producción Animal, Universidad Autónoma Chapingo*, Carretera Federal México-*Texcoco Km 38.5, 56227, Texcoco, Estado de México, México.*

[b] *Socio Economía Estadística e Informática, Colegio de Postgraduados, Carretera Federal México-Texcoco Km 36.5, 56230, Texcoco, Estado de México.*

Corresponding author: Paulino Pérez-Rodríguez, Email: perpdgo@gmail.com

Corresponding author ORCID: 0000-0002-3202-1784

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 animals was considered. Observed heterozygosity (Ho) revealed high genetic diversity for the studied traits, Ho=0.42, 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 genetic diversity that formed at least two groups of animals by genotype. These results coincide with what was observed in the biplot of non-metric multidimensional scaling. Allelic frequencies for the groups formed with the K-means algorithm were estimated as well. It was found that this grouping strategy allowed visualization of distant groups by genotype but not by favorable alleles in all the loci. In addition, the genes *CSNK1E*, *DNAH11*, *DSC2*, *IBSP*, and *OCLN* affected most of the traits and they were highly informative. Therefore, they represent a potential resource for selection 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.

Key words: Brown Swiss, candidate gene, cluster analysis, genetic variability.

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.

Today, the strategy in most dairy and dual-purpose herds has emphasized selection of productive and reproductive traits jointly to make the production system more efficient. Since 2000, genetic and genomic selection has included, among others, characteristics such as cow and heifer conception, and positive production results have

correlated highly with the animals' genetic merit (Veronese et al., 2019). Recently, single nucleotide polymorphisms (SNP) have become a widely used resource for genomic selection. One of the reasons is that the cost of genotyping has decreased to a cost-effective level.

Genetic diversity is a key element in any genetic improvement program. 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 (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 production systems because of its good results for beef production. 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 characterize genetically a Mexican Braunvieh population for productive life, pregnancy rate, and cow and heifer conception rate, through an analysis of genetic diversity

focusing on genotypes and favorable alleles for intra-genic SNPs of candidate genes, using expected and observed heterozygosity indexes, grouping algorithms, analysis of non-metric multidimensional scaling, and estimation of allele frequencies.

Material and methods

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.

Genotype quality control

For data edition before genotype analyses, quality control was performed. 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 to discard non-informative markers or markers with errors in the genotyping.

Statistical analyses

Identification of associated and informative loci

The candidate genes associated with reproductive traits in *Bos taurus* were identified. 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 associated with pregnancy rate (PR), heifer (HCR) and cow conception rate (CCR), and productive life (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 $\log(\cdot)$ denotes the natural logarithm function. This index was calculated for each of the loci. Empirical distribution was obtained, and the 10 or 12 most informative markers were selected. The highest values for H were associated with the highest diversity for the marker.

Cluster analyses

The groups formed by their genetic diversity were generated using hierarchical grouping based on the Euclidian distance matrix as a measure of dissimilarity and the Ward distance as a measure of proximity between groups (Ward, 1963). The analysis was performed using the cluster function included in the stats package of R (R Core

Team, 2021). Additionally, data were also clustered using the K-means algorithm, using the kmeans 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. K-means is a non-hierarchical method of re-assignment that allows grouping similar individuals, the method finds K non-overlapped groups (Wu, 2012). According to Mesquita et al. (2017), the Euclidean distance η_{ij} between two vectors $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$, of dimension D, is given by:

$$\eta_{ij} = z_{ij}^{1/2} = \sqrt{\sum_{d=1}^{D} (x_{i,d} - x_{j,d})^2},$$

where $z_{ij} = \|\boldsymbol{x}_i - \boldsymbol{x}_j\|_2^2 = \sum_{d=1}^{D} (x_{i,d} - x_{j,d})^2$ is the squared distance between $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$, with $i,j = 1,...,n$ and n is the number of individuals. In the context of our analysis, the vectors $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ represent the genotypic information recoded for additive effects for individuals *i* and *j*.

To validate the results found using hierarchical grouping and the K-means method, the data was also analyzed using the Partition Around Medioids (PAM) algorithm (Lengyel and Botta-Dukát, 2019). PAM yields a statistic and a graph which determines grouping quality on a global and individual level using silhouette width and average silhouette width for each group and average silhouette width including all individuals. 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).

Non-metric multidimensional scaling

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 multi-variant analysis was combined with the K-means results to visualize graphically the clusters and the animal with the representative genotype of each group (the closest to the centroid).

Heterozygosity

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 package boot in R (Canty and Ripley, 2021).

Allelic frequencies and proportion test

The allelic frequencies for each group formed with the K-means algorithm were obtained for each studied trait using the software Cervus 3.0.7 (Kalinowski et al., 2007). To determine whether significant differences existed between allelic frequencies for desirable alleles of the groups formed, and those of representative animals, proportion X^2 tests were carried out using the *chisq.test* 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 visualize proportions for the frequencies of favorable alleles of the most informative genes in the database that affect 3 to 4 of the studied traits: *CSNK1E*, *DNAH11*, *DSC2*, *IBSP*, and *OCLN*.

Results and Discussion

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 favorable allele reported in literature is in parentheses. For some of the genes, there are over-dominance reports; therefore, the favorable 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.

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 higher. 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.

Clustering using the Ward and K-means algorithms

The average silhouette width for all the traits obtained with the Ward method was 0.13, while 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 formed 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, relative to the clustering power of the methods for characterizing genetic diversity, coincide with Campos et al. (2017), who used principal components analysis (PCA) 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 favorable 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 (Table 4 and Table 1). 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, while 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 favorable 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

could 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 favorable alleles.

The few differences between proportions are attributed to a compensatory effect, a consequence of analyzing 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 favorable 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 favorable 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 favorable 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 favorable 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.

Non-metric multidimensional scaling

In Figure 2, the bi-dimensional representations obtained with NMDS for the studied traits are shown. To visualize the grouping by genotypes, the K-means classification was included in the graphic, highlighting each of the groups found and the representative animals of 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, while the value for PL (0.15) permits graphic visualization of what NMDS offers (Dexter et al., 2018). NMDS is a useful tool to visualize 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 center 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 neighboring groups.

NMDS is a tool widely used in genetic diversity studies. Like the findings of our study, satisfactory results have been found in the graphic representation of the method, for example, the study of Moscarelli et al. (2020) with breeds derived from the original Braunvieh and the study of Senczuk et al. (2020) where NMDS was used to visualize bi-dimensionally differences among cattle breeds from the Alpine arc.

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 characterize 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, checked out using the *pwr.p.test* function from the *pwr* package (Champely, 2020), was sufficient to detect differences between groups. 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 favorable 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 analyzed 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 favorable 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 favorable alleles for *DSC2* and *LHCGR*, while group I had the highest number for *OCLN*.

Group I had the lowest frequency of favorable 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 favorable 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*, *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 favorable 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*), while 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 generalized 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 favorable 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 favorable 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 favorable 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 favorable 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.

Final considerations

The results of our study suggest the presence of genetic diversity for the studied traits in the population analyzed. 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 characterization 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.

Ethics approval

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

Author contributions

MZTC drafted the first version of the manuscript, PPR supported MZTC during the analysis, ARF and RLO conceived the project. All the authors contributed to obtain the final version, PPR acted as the author for correspondence.

Conflict of interest

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

Financial support

This work was in part supported by a grant from the Universidad Autónoma Chapingo, México (DGIP-166701012).

Data availability

Upon reasonable request to the author for correspondence.

Author ORCIDs

Mitzilin Zeleica Trujano-Chavez 0000-0002-3242-1040

Rufino López-Ordaz 0000-0002-1825-0097

Agustín Ruíz-Flores 0000-0001-8267-2107

Paulino Pérez-Rodríguez 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. doi: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. doi:10.1186/s12864-020-6446-y

Campos, B. M., A. S. do Carmo, A. A. do Egito, A. S. da Mariante, M. S. M. do Albuquerque, S. J. J. de Gouveia, M. C. H. Malhado, L. L. Verardo, B. M. V. G. da Silva, S. P. L. & Carneiro., 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. doi: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. doi: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. doi: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. doi: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.

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. doi: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. doi: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.

Mesquita, D. P. P., Gomes, J. P. P., Souza Junior, A. H., & Nobre, J. S., 2017. Euclidean distance estimation in incomplete datasets. *Neurocomputing,* 248, 11–18. doi:10.1016/j.neucom.2016.12.081

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). doi: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. doi: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

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. doi: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. doi:10.1111/age.12420

Peixoto, M. G. C. D., Carvalho, M. R. S., Egito, A. A.; Steinberg, R. S. Bruneli, F. Â.T., 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. doi: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/

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. doi: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(40), 1-12. doi:10.1186/s12711-020-00559-1

Sherwin, W. B., 2010. Entropy and information approaches to genetic diversity and its expression: Genomic geography. *Entropy,* 12 (suppl. 7), 1765–1798. doi: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. doi: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. doi: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. doi: 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). doi: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. doi:10.3168/jds.2018-15205.

Ward, J. H., 1963. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association,* 58 (suppl. 301), 236–244. doi:10.1080/01621459.1963.10500845

Wu, J., 2012. Advances in K-means Clustering. Springer Theses, PhD thesis, Tsinghua University, Pekín, China. doi:10.1007/978-3-642-29807-3

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,* doi: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. doi:10.1007/s11250-018-1655-y

564 **Table 1.** Top 12 and 10 single nucleotide polymorphisms (SNP) associated with

565 pregnancy rate and cow productive life in Braunvieh Cattle.

Pregnancy rate		Productive life	
Gene (DA)[1]	**SNP**	**Gene (DA)[1]**	**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		

566 Gene (DA) = Gene associated with the reproductive trait, in parenthesis the desirable allele reported by

567 Cochran et al. (2013) and Ortega et al. (2015). Gene°, genes that have been reported with overdominance

568 or dominance.

Table 2. Top 10 single nucleotide polymorphisms (SNP) associated with heifer and cow conception rate in Braunvieh Cattle.

Cow conception rate		Heifer conception rate	
Gene (DA)[1]	**SNP**	**Gene (DA)[1]**	**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.

574 **Table 3.** Genetic diversity indices by marker associated with reproductive traits in the
575 Mexican Braunvieh cattle population: observed (Ho) and expected (He) heterozygosity
576 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

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.

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.

Genes	PRG		PLG			CCRG		HCRG	
	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

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.

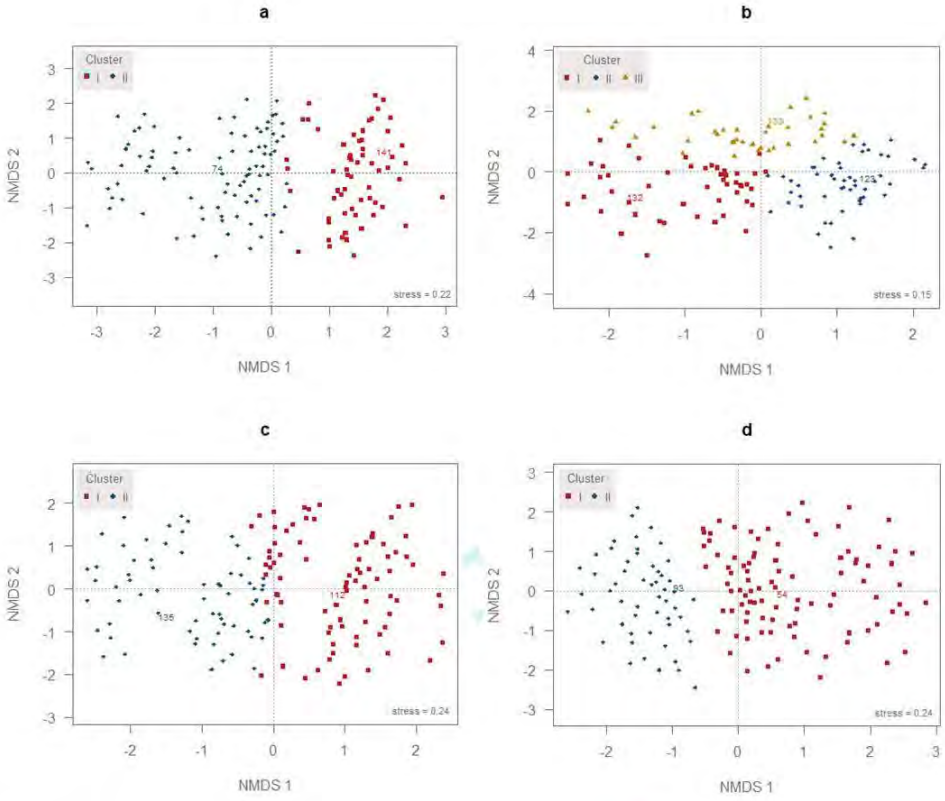

Figure 2. Graphic representation of the non-metric multidimensional scaling (NMDS) for Pregnancy Rate (a), Productive Life (b), Cow Conception Rate (c), and Heifer Conception Rate (d), using K-means clusters to visualize graphic differences and the representative animals, in a Mexican Braunvieh Cattle population.

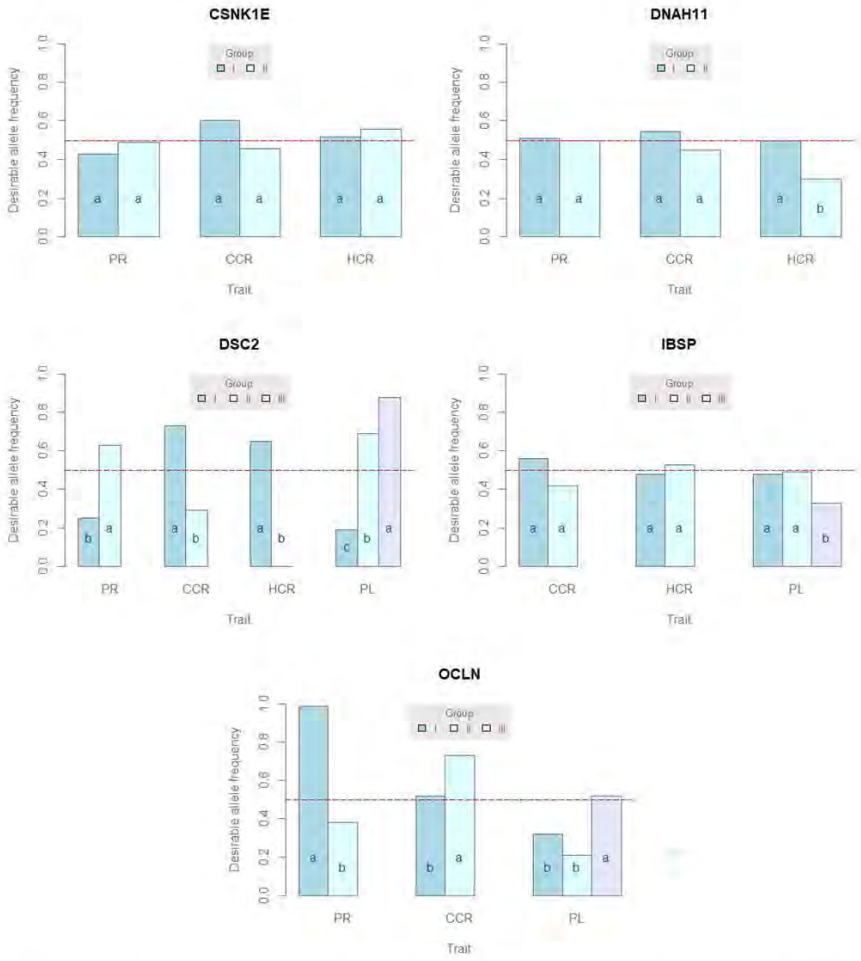

Figure 3. Graphic representation of proportion tests using X^2 test between K-means on groups classified by reproductive trait for frequencies of desirable alleles of the most important genes in a Mexican Braunvieh population. PR = Pregnancy rate. PL = Productive life. CCR = Cow conception rate. HCR = 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.

CAPÍTULO 5. GENETIC DIVERSITY AND POPULATION STRUCTURE FOR RESISTANCE AND SUSCEPTIBILITY TO MASTITIS IN BRAUNVIEH CATTLE

veterinary sciences

MDPI

Article

Genetic diversity and population structure for resistance and susceptibility to mastitis in Braunvieh cattle

Mitzilin Zuleica Trujano-Chavez [1], Reyna Sánchez-Ramos [2], Paulino Pérez-Rodríguez[3] and Agustín Ruíz-Flores [1,*]

[1] Posgrado en Producción Animal, Universidad Autónoma Chapingo, Carretera Federal México-Texcoco Km 38.5, 56227, Texcoco, Estado de México, México; zulealizee@gmail.com; aruizf@chapingo.mx
[2] Recursos Genéticos y Productividad; rey_1014@hotmail.com
[3] Socio Economía Estadística e Informática-Estadística, Colegio de Postgraduados, Carretera Federal México-Texcoco Km 36.5, 56230, Texcoco, Estado de México.; perpdgo@gmail.com
* Correspondence: aruizf@chapingo.mx; Tel.: +52 595 952 1621

Abstract: Mastitis is a disease that causes significant economic losses since resistance to mastitis is a difficult trait to improve due to its multifactorial occurrence. Therefore, our objective was to characterize a Mexican Braunvieh cattle population for genetic resistance and susceptibility to mastitis. We used 66 SNP markers for 45 candidate genes in 150 animals. Average heterozygosity was 0.445±0.076, a value higher than those reported for some European breeds. The inbreeding coefficient was slightly negative for resistance to subclinical (-0.058±0.055) and clinical (-0.034±0.076) mastitis, possibly due to low selection for the immunological candidate genes that influence these traits. The genotypic profiles for the candidate loci per K-means group were obtained, as well as the groups distribution through the graphics of the principal component analysis. The genotypic profiles showed high genetic diversity among groups. Resistance to clinical mastitis had the lowest presence of heterozygous genotypes. Although the percentage of highly inbred animals (>50%) is up to 13.3%, there are highly heterozygous groups in terms of the studied traits, a favorable indicator of the presence of genetic diversity. The results of this study constitute evidence of the genetic potential of the Mexican Braunvieh population to improve mastitis-related traits.

Keywords: inbreeding, heterozygosity, candidate genes, principal components analysis

Citation: Citation: Trujano-Chavez, M.Z.; Sánchez-Ramos, R.; Pérez-Rodríguez, P.; Ruiz-Flores, A. Genetic diversity and population structure for resistance and susceptibility to mastitis in Braunvieh cattle. *Vet. Sci.* 2021, *8*, x. https://doi.org/10.3390/xxxxx

Academic Editor: Firstname Lastname

Received: date
Accepted: date
Published: date

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

One of the major problems in dairy units is mastitis [1]. Tropical dual-purpose production systems have the same problem. A prevalence of 60% was reported for mastitis in tropical production units [2]. The current trend seems to focus on milk quality rather than on quantity, and mastitis has been one of the main hurdles [1,3].

Some of the obstacles that traditional genetic selection has faced in bovines are [4]: 1) complexity of quantitative or low-heritability traits, 2) traits measured late in the life of the animals, and 3) traits that are difficult to measure, such as resistance and susceptibility to mastitis.

Resistance to mastitis is a complex trait that is multifactorial in nature and, thus, difficult to improve. The genomic prediction ability for clinical mastitis is low, 0.19 [5], compared to that reported for residual feed intake (0.85±0.08), dry matter intake (0.52±0.04) and feed conversion ratio (0.46±0.05) [6]. The conventional selection scheme for resistance has not been effective against mastitis due to the large number of candidate genes related to several non-immune factors that influence it and challenge its measurement [4, 7-8].

In addition to the control of environmental factors to prevent it at the farm level, genomic methods have been used as an alternative solution to the problem of increasing resistance to mastitis. The aim is to obtain animals that are immune to the main mastitis-

causing pathogens and have qualities that allow them to adapt to environmental factors that cause mastitis [2].

Recently, genotyping of animals has become easy and relatively cheap. A sample of hair or blood will provide thousands of markers [4]. The markers most used for genomic selection are the single polymorphism nucleotides (SNP). Their generalized use has reduced the cost of genotyping to a cost-effective level. Genotyping is widespread, used in both commercial and research systems [9].

Manipulation of genetic variation is key to the development and use of any genetic improvement program for dairy and beef cattle breeds [10]. Genetic diversity has been studied in original Braunvieh and derivative breeds with the aims of disentangling the genetic basis of their adaptation to diverse environments and increasing productive performance. The results have been positive. It was found that populations of this breed adapted to environments different from that of its origin are genetically more diverse than the average of other *Bos taurus* and some Bos indicus breeds [11]. Animals living in harsh environmental conditions, such as extreme temperatures, possess greater genetic diversity because facing environmental changes requires versatility [12].

The objective of this study was to determine the genetic diversity and structure of a Mexican Braunvieh population in terms of mastitis resistance and susceptibility, using 1) basic population statistics: inbreeding coefficient, expected and observed heterozygosity, and 2) multivariate statistics: K-means and principal components analysis.

2. Materials and Methods

2.1. Source of information

Genomic information was obtained using 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

For data edition, quality control was performed prior to the genotypic analyses. 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 entire population. Additionally, a 0.05 threshold for the minor allele frequency was considered to discard non-informative markers or those with errors in genotyping.

2.3. Statistical analyses

2.3.1. Identification of associated and informative loci

The candidate genes associated with mastitis in *Bos taurus* were identified, according to reports by [13-16] for resistance to mastitis (RM); by [17-18] for susceptibility to mastitis (SM); by [19-21] for resistance to bacterial mastitis (RBM); by [8] for resistance to subclinical mastitis (RSM); and by [22] for resistance to clinical mastitis (RCM).

The position ranked by base pairs were obtained from the Gene Library [23]. With this information, the intragenic loci were searched; the criterion for the loci-gene association was that they should be located within the position rank ± 25 kb of the gene for the same chromosome. Only the loci available in the genotypic database were included. The analyses considered 455 SNP markers of 204 candidate genes reported in the literature associated with resistance and susceptibility to mastitis.

All the analyses were carried out using the R software, version 4.0.4 [24]. According to the literature consulted, the 10% most informative markers associated with each of the

studied traits were determined using the Shannon index. This index allows identification of the most informative markers [25]. 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. This index was calculated for each of the loci. Empirical distribution was obtained, and the 10% most informative markers were selected. The highest values for H were associated with the highest diversity for the marker. At the end of filtering, 66 SNP markers of 45 candidate genes were left for further analyses. For better handling of the markers in the graphs, the names of the loci were replaced by consecutive numbers, assigned by alphabetical order to the corresponding candidate gene.

2.3.2. Population genetic structure

Expected (He) and observed (Ho) heterozygosity was estimated for all the loci by trait in the study by using the adegenet program [26] available in R [24]. The stats program of R [24] was used to perform the t-test to determine the differences between He and Ho. The null hypothesis was that there were non-significant differences between them. The p-value 0.01, used to declare significance, and the p-value was corrected with the Bonferroni method considering the total number of loci by trait.

The program pegas [27] through adegenet [26] was used to determine whether the loci under study were in Hardy-Weinberg equilibrium. A hundred thousand replicates were assigned to execute the test procedure through the Monte Carlo method. The null hypothesis established the existence of Hardy-Weinberg equilibrium when p-values were higher than $0.01/n$ with the Bonferroni correction; n is the number of loci per trait.

The inbreeding coefficients by individual with respect to the population (F_{IT}) were calculated and plotted with adegenet [26]. Additionally, the general mean for F_{IT} per trait was obtained, along with its respective standard deviation. Finally, a histogram including a nonparametric estimate of the density function by trait was obtained. This was done to visualize the animals with an inbreeding coefficient above 0.5 (the number of animals could be observed in the histogram).

2.3.3. Cluster analyses

The differences in the genotypic profile of individuals were identified with two clustering algorithms: the hierarchical clustering using the Ward distance implemented in the stats package and the K-means algorithm also included in the stats package in R and factoextra [28]. For both methods, the 10% of the markers, identified in a previous step, was used.

The optimum number of groups (k) to be formed through the hierarchical clustering and K-means was obtained with Silhouette and Elbow of factoextra [28]. A search for k that allowed equilibrium between a value of the silhouette close to one [29] and the least error sum of squares was carried out.

The Ward algorithm of minimum variance, according to [30], is a divisive hierarchical method that creates mutually exclusive sub-sets. Sub-set members are ultimately similar to each other and different between groups. This promises to be a good method for finding different groups with particular genotypic profiles. The results of the procedure are illustrated in a circular dendrogram that displays the hierarchy of the groups and their similarity to the groups formed with the K-means algorithm.

K-means is a non-hierarchical algorithm that uses a reassignment method based on centroids to form groups. This is the most frequently used method to form groups [31]. It is used in robotics, genomics, and genetic diversity. In our study the groups of K-means were used to identify divergent genotypes among groups of animals and to characterize genetic diversity based on the Principal Components Analysis (PCA).

2.3.4. Principal Components Analysis

The PCA was performed using adegenet package [26] using the sample covariance matrix obtained using the candidate SNP. The objective was to observe the distribution of the groups formed by K-means through a PCA graphic of individuals (1st and 2nd dimensions) and to reduce the number of markers according to their eigenvalues. With a factoextra biplot [28], we estimated the correlation and contribution of markers to the 1st and 2nd dimensions formed by the PCA and determined the 10% of the animals with the highest contribution to the analysis.

3. Results and Discussion

3.1. Sample size and power analysis

In México, studies using genomic information from large animal populations are scarce. Usually, the main restriction is the availability of economic resources for animal genotyping. Genetic improvement infrastructure of developing countries usually is less robust than that of developed countries whose genomic databases involve millions of animals [32].

In previous studies carried out with the same population used in our study [33-34], genomic information of 300 animals was considered. For our study only 150 animals with candidate SNPs associated with resistance and susceptibility to mastitis were found.

The sample size used in our study, evaluated with the pwr.chisq.test function of the pwr package [35], was enough to detect significant differences with a p-value of 0.05 and an effective size of 0.4. The power of 0.87 for the Xi^2 test obtained in the present study is adequate to obtain conclusive, reliable results [36]. In other studies of genetic diversity and structure, smaller sample sizes have been used, and conclusive results have been reached: 6 to 23 [37], 7 to 58 [38], 13 to 38 [39], and 71 to 167 [40].

3.2. Identification of associated and informative loci

Tables 1 to 5 show the markers obtained after filtering with the Shannon Index. Their candidate genes, chromosome, and the number assigned to visualize the results in the graphs are shown. For RM, 19 genes with 22 loci were obtained; for SM, seven genes and 11 loci; for RBM, seven genes and eight markers; for RSM, eight genes and 20 markers; and for RCM, four genes and six markers. This adds up to 66 markers and 45 candidate genes. All loci were in Hardy-Weinberg equilibrium.

3.3. Population genetic structure

Table 6 shows the results for He and Ho, as well as those of a t-test to determine whether the differences were statistically significant. The rule of decision to not reject the null hypothesis of non-significant differences between He and Ho was that the p-value for the t-test > p-value for BC. In the last column the |He-Ho| differences can be observed; all were non-significant, with exception of the difference for SM.

The estimates for Ho obtained in our study for mastitis (average 0.45±0.076) are above those obtained with breeds derived from the original Braunvieh, but below those reported for *Bos taurus* breeds with some adaptation to the tropics. In Colombia for Creole cattle, [41] reported Ho of 0.66; the Costeño con Cuernos breed had the lowest, 0.635, and Casareño the highest, 0.733. These Ho estimates are quite different from those reported by [42], who found an average of 0.35±0.167 for Ho in populations related to Braunvieh.

This suggests that the Mexican Braunvieh population has increased its genetic versatility for the candidate genes studied, with a higher number of heterozygous individuals. Exposure to climates different from their native climate for over a century in the tropical systems of Mexico could have caused a change in Ho, with respect to original Braunvieh populations [42-43].

Table 1. Top 10% single nucleotide polymorphisms (SNP) associated with Resistance to Mastitis (RM) in Braunvieh Cattle.

Gene	Chr	Position, pb (± 25 k)	SNP marker name	LN
ARHGAP10 [15]	17	9994895-10429767	ARS-BFGL-NGS-113821	1
			BovineHD1700002890	2
BDH2 [14]	6	21737056-21815795	BovineHD0600006001	3
CAPG [13]	11	49562863-49630835	Hapmap54495-rs29018810	4
CHD5 [13]	16	47131943-47253618	BovineHD1600013051	5
CST6 [14]	29	44076083-44127573	BovineHD2900013196	6
ELMO1 [13]	4	59842686-60475511	BovineHD0400016236	7
FBL [15]	18	49363469-49425788	ARS-BFGL-NGS-113564	8
FOCAD [13]	8	23394437-23750228	BovineHD0800007122	9
IMMP2L [13]	4	56736178-57747698	BovineHD0400015680	10
LOC510112 [15]	10	27914007-27964951	BovineHD1000009188	11
			BovineHD1000009179	12
MBL2 [16]	26	6306933-6362539	BovineHD2600001441	13
			BovineHD2600001446	14
MYO1E [13]	10	50932729-51205381	BovineHD1000015285	15
PON1 [13]	4	12517349-12601241	ARS-BFGL-NGS-48351	16
PPP3CA [13]	6	23419690-23795287	BovineHD0600006555	17
ST7 [13]	4	51162819-51481394	BovineHD0400014185	18
SULF2 [14]	13	76130676-76307180	BovineHD1300022072	19
TBC1D8 [13]	11	5953170-6144340	BovineHD1100002208	20
TBCK [13]	6	18840254-19099244	BovineHD0600005226	21
ZNFX1 [14]	13	77323734-77400021	BovineHD1300022388	22

Chr = Chromosome. Position, pb (± 25 k) = Candidate gene position in base pairs reported in [23], for *Bos taurus* ± 25,000 base pairs. LN = Locus Number assigned in graphs.

On the other hand, the non-significant differences between He and Ho for most of the traits (Table 6) is similar to the results obtained in other studies, where neither the breed nor the environment modified the non-rejection of the null hypothesis, He=Ho for Xi^2 [41-44].

As a graphic resource, Figure 1 shows that the absolute differences at the locus level are small, as in the overall result. The exception to this was SM whose joint trend of the loci was to be over 0.03. Thus, it is the only trait with significant differences. In the barplots for the rest of the traits, there were loci that were both close to zero and above 0.05, but the overall estimate is not enough to declare evidence for a significant difference.

The estimate for Fit obtained coincides with those observed in semi-specialized systems of production. The overall average for Fit was 0.017±0.043 for all the traits associated with mastitis. This value is similar to those found for Sahiwal (0.013±0.109), Gyr (0.013±0.106), and Guernsey (0.02±0.217) [45]. In contrast, the Fit for the tropical breeds Landim, Angole, and Tete [46] was 70% higher than that found in our study. Furthermore, for Creole Colombian breeds the average was 0.09 [41]. The first group of breeds was maintained in semi-specialized systems [45], while Landim, Angole, Tete, and the Colombian Creole are non-specialized breeds. This difference could explain why smaller breeds without adequate genetic control present higher inbreeding coefficients.

Table 2. Top 10% single nucleotide polymorphisms (SNP) associated with Susceptibility to Mastitis (SM) in Braunvieh Cattle.

Gene	Chr	Position, pb (± 25 k)	SNP marker name	LN
ANKRD33B [17]	20	62576901-62728565	ARS-BFGL-NGS-112060	1
CTNND2 [17]	20	61189491-62339843	BovineHD2000017160	2
			BovineHD2000017153	3
			BTB-00791947	4
GRIA3 [18]	10	7542786-7901396	ARS-BFGL-NGS-87466	5
ILDR2 [18]	3	1878471-1997329	Hapmap42630-BTA-67480	6
ITPK1 [18]	21	57564349-57778522	BovineHD2100016547	7
			BovineHD2100016592	8
			BovineHD2100016552	9
SIDT1 [18]	1	58193233-58355060	BovineHD0100016494	10
TBXAS1 [18]	4	103270503-103494001	BovineHD0400029111	11

Chr = Chromosome. Position, pb (± 25 k) = Candidate gene position in base pairs reported in [23] for *Bos taurus* ± 25,000 base pairs. LN = Locus Number assigned in graphs.

Table 3. Top 10% single nucleotide polymorphisms (SNP) associated with Resistance to Bacterial Mastitis (RBM) in Braunvieh Cattle.

Gene	Chr	Position, pb (± 25 k)	SNP marker name	LN
ECHDC1[1] [21]	9	23922553-24025911	BovineHD0900006460	1
GNAI1[2] [19]	4	40930829-41086584	ARS-BFGL-NGS-110306	2
IL1A[3] [20]	11	46457553-46518533	BovineHD1100013546	3
IL1RN[2, 3] [19]	11	46790914-46862407	ARS-BFGL-NGS-113289	4
	11	46790914-46862408	ARS-BFGL-NGS-113289	4
ITGA4[2] [19]	2	15058580-15200364	BovineHD0200004269	5
ITGB3[2] [19]	19	46305531-46418474	UA-IFASA-8333	6
NRG1[3] [21]	27	28396076-28680095	BovineHD2700007962	7

Superscripts indicate resistance to [1] *S. aureus*, [2] *S. agalactiae* and [3] *E. coli*. Chr = Chromosome. Position, pb (± 25 k) = Candidate gene position in base pairs reported in [23] for *Bos taurus* ± 25,000 base pairs. LN = Locus Number assigned in graphs.

Figure 2 shows that, for traits RM, SM, and RBM, most of the loci had positive Fit. In contrast, the loci for RSM and RCM are mostly negative; that is, there are no traces of inbreeding. This could be due to the specificity of the last two traits, where the candidate genes involve DNA segments that have undergone null direct selection.

The genes whose symbol starts with CXCL are responsible for the immune response [47]. In our study, for RCM, 67% of the SNP markers belong to the genes CXCL1 and CXCL8. This explains the low Fit values for the trait, This is because current improvement programs for the Braunvieh population under study have no selection criteria based on resistance to mastitis.

For RSM, 40% of the markers belong to EDN2 and HIVEP3, candidate genes associated with immune processes, such as intracellular and sequential signaling of immunoglobulin receptors. Because of their immunologic nature, these genes have not been directly selected for; however, at the farm level there is a rising concern for improving their animals for genes associated with diseases that could be prevented, such as paratuberculosis and clinic mastitis [8, 48].

Table 4. Top 10% single nucleotide polymorphisms (SNP) associated with Resistance to Subclinic Mastitis (RSM) in Braunvieh Cattle.

Gene	Chr	Position, pb (± 25 k)	SNP marker name	LN
BMPR1B	6	29346363-29845366	BovineHD0600008191	1
			BovineHD0600008283	2
			BovineHD0600008292	3
EDN2	3	104654329-104730954	BovineHD0300029984	4
			BovineHD0300029996	5
GUCA2A	3	103990493-104056213	BTA-98478-no-rs	6
HEYL	3	106417713-106483224	BTB-00148619	7
HIVEP3	3	104084087-104699365	ARS-BFGL-NGS-35125	8
			BovineHD0300029904	9
			BovineHD0300029925	10
			BovineHD0300029955	11
			BTB-01612988	12
			Hapmap48983-BTA-100103	13
MACF1	3	106564750-106955676	ARS-BFGL-NGS-38199	14
			BovineHD0300030641	15
			BovineHD0300030658	16
			BovineHD0300030677	17
			BovineHD0300030702	18
MFSD2A	3	106097951-106160784	BovineHD0300030435	19
SH3PXD2A	26	24136301-24435553	BovineHD2600006239	20

Chr = Chromosome. Position, pb (± 25 k) = Candidate gene position in base pairs reported in [23] for *Bos taurus* ± 25,000 base pairs. LN = Locus Number assigned in graphs. Candidate genes reported by [8].

Table 5. Top 10% single nucleotide polymorphisms (SNP) associated with Resistance to Clinical Mastitis (RCM) in Braunvieh Cattle.

Gene	Chr	Position, pb (± 25 k)	SNP marker name	LN
CXCL1	6	89047989-89100128	BovineHD0600024410	1
CXCL8	6	88785817-88839572	ARS-BFGL-NGS-17376	2
			BovineHD0600024315	3
			BovineHD0600024328	4
SEL1L	10	92733415-92848117	BovineHD1000026808	5
STAT4	2	79543834-79730325	BovineHD0200022927	6

Chr = Chromosome. Position, pb (± 25 k) = Candidate gene position in base pairs reported in [23] for *Bos taurus* ± 25,000 base pairs. LN = Locus Number assigned in graphs. Candidate genes reported by [22].

The distribution of animals by their Fr value can be observed in Figure 3. In general, the animals fall in the range of 0 to 0.3 Fr. Fr=50% is equivalent to the value of an inbred animal produced by two progenitors with 100% genetic relationship. Our animals varied from 1.3% to 13.3%. The trait with the highest number of highly inbred animals was SM, while RSM had the lowest.

Although up to 13.3% of animals have non-desirable Fr values, the estimate obtained in our study (0.017±0.043) is similar to that reported for other cattle breeds in specialized

and semi-specialized systems of production. In those systems, although some animals had high inbreeding coefficients, the average F_{IT} is lower than those reported by [49] for *Bos taurus* breeds (0.071), Brown Swiss (0.071), Braunvieh (0.059), Original Braunvieh (0.023), Holstein (0.057), and Simmental (0.028).

Table 6. Heterozygosity and t-test results by trait associated with mastitis, for a Mexican Braunvieh Cattle population.

Trait	He ± SD	Ho ± SD	NL	p-value BC	p-value t-test	He-Ho
Resistance to Mastitis	0.495 ± 0.006	0.468 ± 0.034	22	4.5×10^{-4}	5.5×10^{-4}	0.027
Susceptibility to Mastitis	0.495 ± 0.004	0.458 ± 0.019	11	9.1×10^{-4}	1.0×10^{-5}	0.037*
Resistance to Bacterial Mastitis	0.488 ± 0.010	0.470 ± 0.031	7	1.4×10^{-3}	4.4×10^{-2}	0.018
Resistance to Subclinical Mastitis	0.379 ± 0.153	0.403 ± 0.164	20	5.0×10^{-4}	9.9×10^{-1}	- 0.024
Resistance to Clinical Mastitis	0.406 ± 0.105	0.426 ± 0.130	6	1.7×10^{-3}	8.9×10^{-1}	- 0.020

He ± SD = Heterozygosity expected ± Standard Deviation. Ho ± SD = Heterozygosity observed ± Standard Deviation. NL = Number of loci. p-value BC = p-value with Bonferroni correction (0.01/NL). p-value t-test = p-value obtained by the t-test, where the null hypothesis established that the differences between He and Ho from zero were non-significant, at 0.01/NL significance level. He-Ho = Differences between heterozygosity values (expected - observed), values* indicate statistically significant difference.

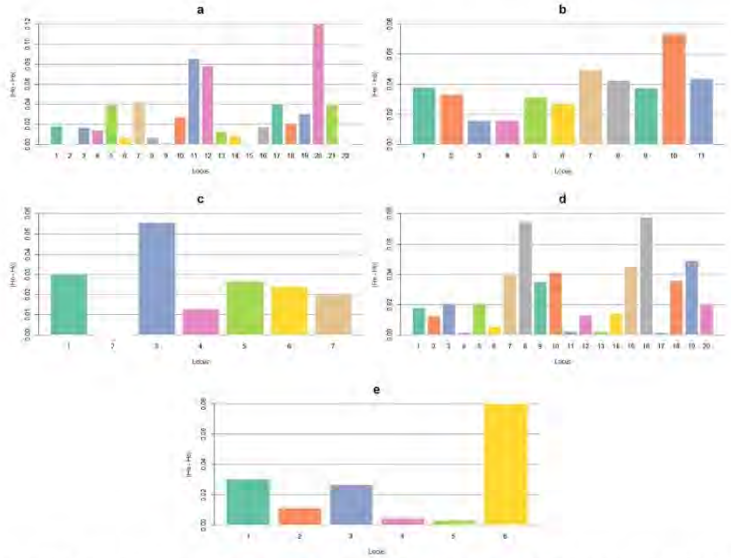

Figure 1. Absolute differences by locus of expected heterozygosity (He) minus observed (Ho) for resistance (a) and susceptibility (b) to mastitis, resistance to bacterial mastitis (c), subclinical mastitis (d) and clinical mastitis (e) in a Mexican Braunvieh cattle population. |He-Ho| = absolute difference between He and Ho. Note: see the names of each locus by trait in Tables 1 to 5.

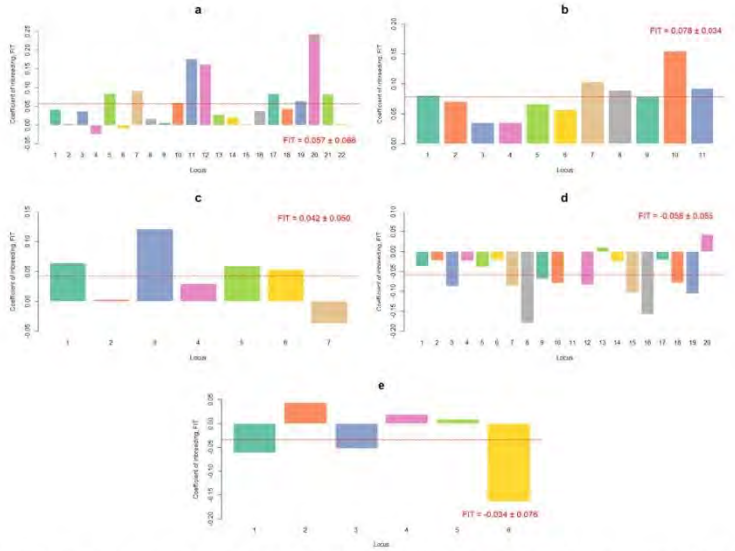

Figure 2. Coefficients of inbreeding (F_{IT}) per locus and trait for resistance (a) and susceptibility (b) to mastitis, resistance to bacterial mastitis (c), subclinical mastitis (d) and clinical mastitis (e) in a Mexican Braunvieh cattle population. The mean of F_{IT} and its standard deviation are shown in red. The red dotted line indicates the mean of F_{IT} for all the loci. Note: see the names of each locus by trait in Tables 1 to 5.

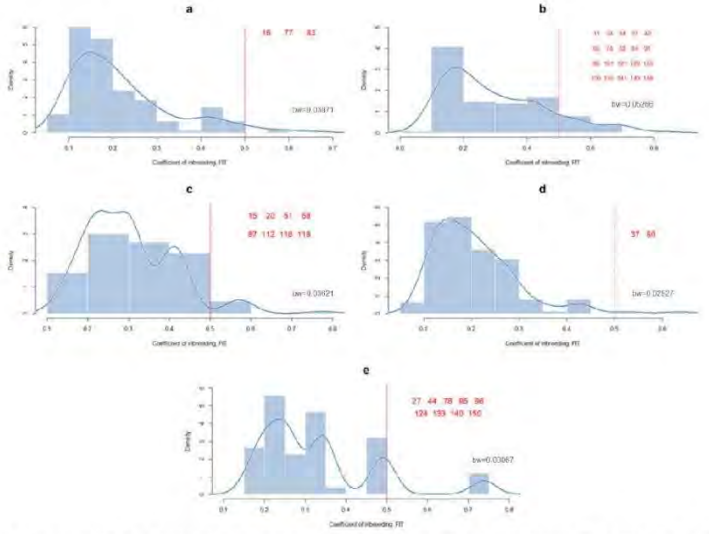

Figure 3. Histogram and density plot (cosine Kernel Function) of the coefficient of inbreeding (F_{IT}) for resistance (a) and susceptibility (b) to mastitis, resistance to bacterial mastitis (c), subclinical mastitis (d) and clinical mastitis (e) in a Mexican Braunvieh cattle population. bw= Bandwidth. Animals with a F_{IT} greater than 0.5 are shown in red on each plot.

3.4. Clustering using hierarchical methods and K-means algorithms

The optimum number of k to equilibrate the silhouette width and the means of squared errors for both algorithms converged at the same number of clusters for each trait. However, better values for silhouette were found by grouping with the K-means algorithm: RM, 0.08; RS, 0.17; RBM, 0.17; RSM, 0.12; and RCM, 0.39. For the first four traits, the silhouette values were low, but in all cases, they were negative, a good indicator of the suitability of the method. The RCM silhouette was the widest.

The average proportion of variability among groups relative to the total variability was obtained. The estimates found were RM, 34.9%; RS, 65.5%; RBM, 66.9%; RSM, 54.7%; and RCM, 77.3%. It should be noted that the RCM values for variability and silhouette were the best of all the traits. This could be due to the small number of markers of the trait (6), which were highly informative as well.

The circular dendrograms of Figure 4 were generated to visualize the groups formed by the Ward hierarchical method. The groups created with this algorithm were very similar to those created with K-means. The k number, the number of individuals per group, and the representative animals together with their genotypic profile can be seen in Figure 5. The representative animal is the animal closest to the centroid of its cluster, and its genotypic profile was obtained to visualize the main differences among group patterns.

Ho and F_{IT} are not related because, unlike inbreeding, heterozygosity is not correlated with the markers [50], even though Ho values close to 0.50 were found for all the traits. In the genotypic profiles, it can be observed that there are groups of animals with low frequency of genotype AB, Figure 5. The traits with the lowest values for F_{IT} present high numbers of homozygous animals. This could be explained by the low genetic relationship among the animals for these loci, that is, genotypes AA or BB for the same locus, as seen in Table 6. One of these traits (SM) presents a significant He-Ho difference. This means a decrease in the expected value for heterozygosity which will be reflected in a greater number of homozygous individuals.

The trait with the greatest number of heterozygous animals was RM. Groups III, VII, VIII, and XI possess most of their loci with genotype AB. RM is multifactorial, and therefore, it includes candidate genes associated with other productive, reproductive, and adaptative traits. This will cause high genetic diversity, in contrast with the traits RSM and RCM, which are influenced by immunological candidate genes of specific order whose diversity is limited. According to [51], these gene types tend to be found in homozygous genotypes in wildlife, given that AB genotypes are associated with proteinic abnormalities that increase their susceptibility to pathogenic diseases.

For SM there are well defined groups of animals whose loci are mostly heterozygous. Groups II, IV, V, XV, and XVII have at least 55% of their loci with genotype AB. The groups with fewest heterozygous individuals were III, XVIII, XIX, and XX with up to 18% AB genotypes. There is high genetic diversity in the population for the trait as was observed for RM; the reason could be the presence of genes influencing traits that are indirectly related with mastitis, and not just loci affecting the immune system [47,51].

The loci of the groups formed by the algorithm K-means for RBM are from 29% (groups II, VI, IX, and XX) to 57% (groups I, III, VIII, and XII) heterozygous. Diversity was high because only 33% of the genes were directly related with immune functions in the animal, genes IL1A and IL1RN, while the rest was related to metabolic processes [19-21].

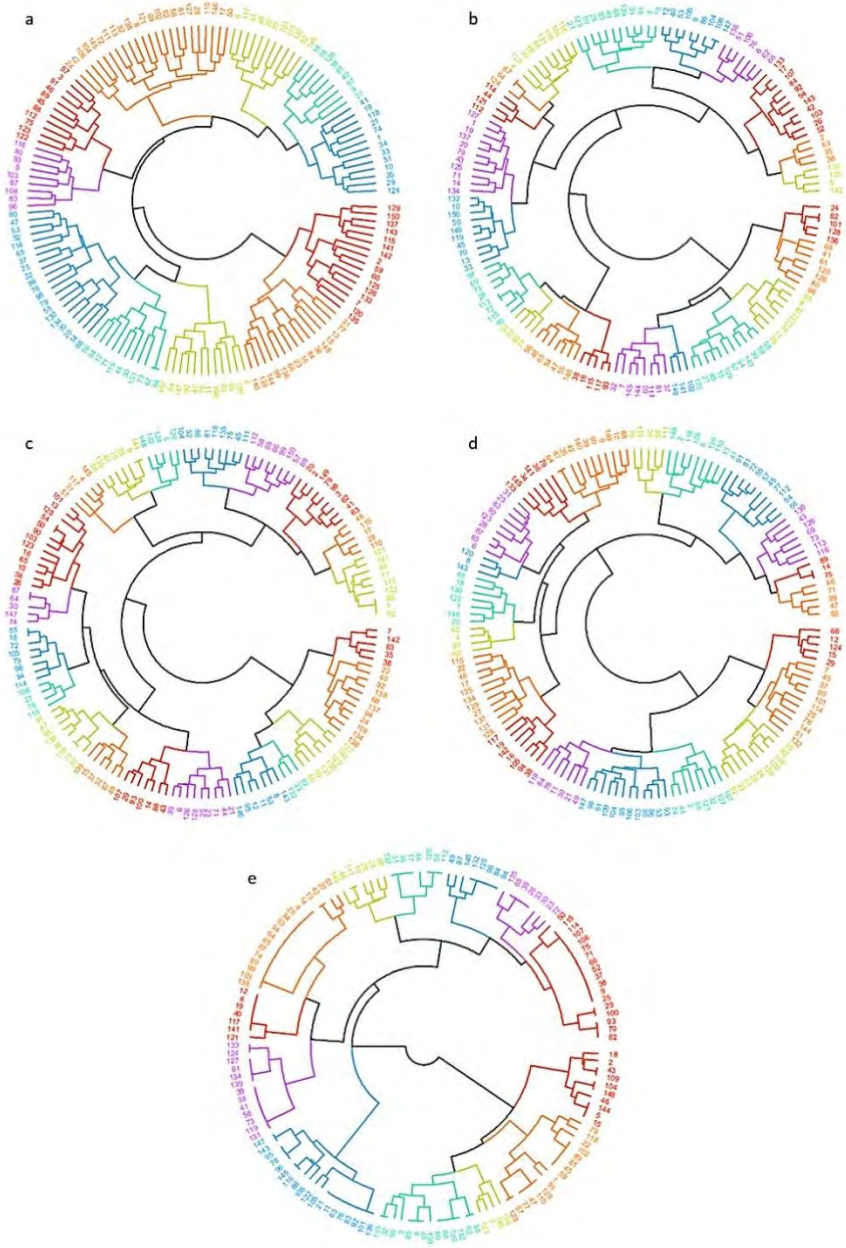

Figure 4. Circular dendrograms obtained by Ward's algorithm based on Euclidian distance for resistance (a) and susceptibility (b) to mastitis, resistance to bacterial mastitis (c), subclinical mastitis (d) and clinical mastitis (e) in a Mexican Braunvieh cattle population.

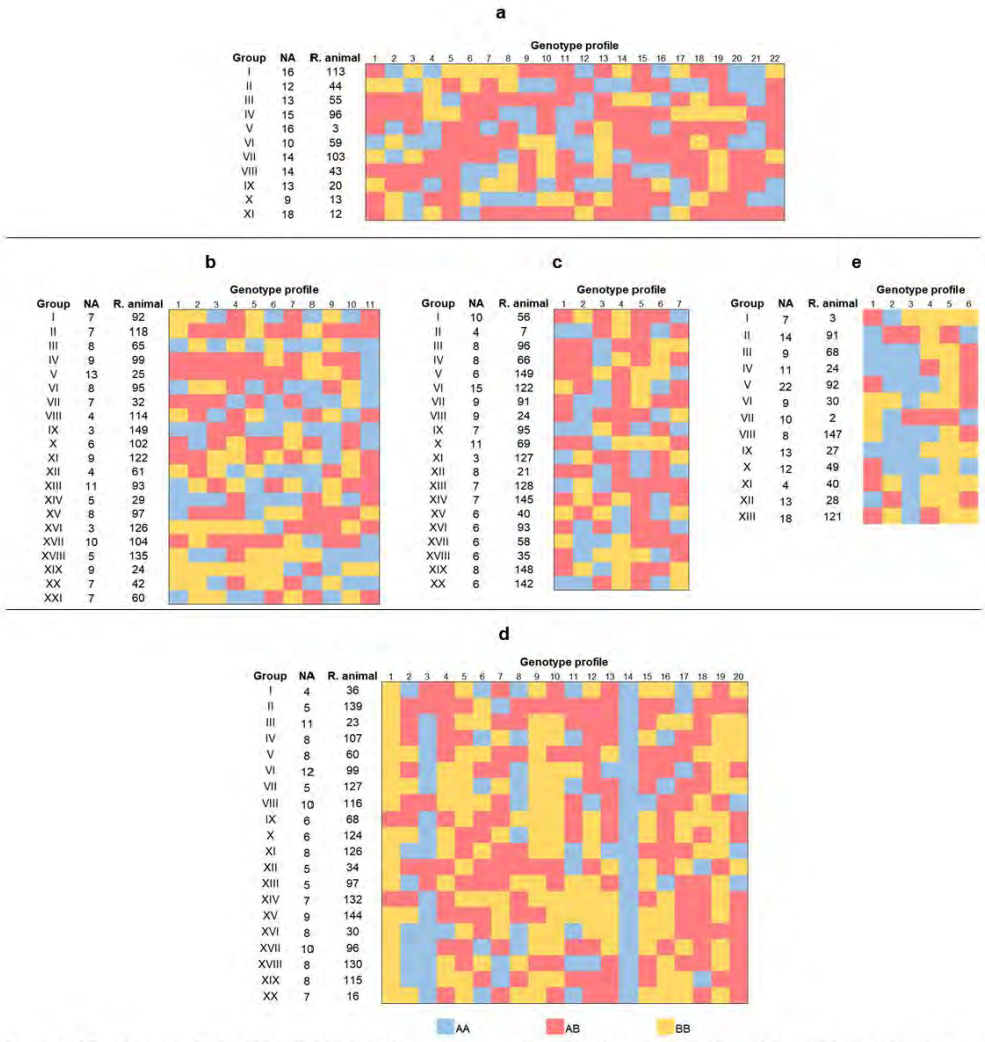

Figure 5. Clustering results for 150 individuals (only one representative animal per group is shown) for resistance (a) and susceptibility (b) to mastitis, resistance to bacterial mastitis (c), subclinical mastitis (d) and clinical mastitis (e) in a Mexican Braunvieh cattle population. NA = Number of animals in the group. R. animal = representative animal (the closest to the centroid). Note: the names of each locus by trait are indicated in Tables 1 to 5.

3.5. Principal Components Analysis (PCA)

The PCA was useful to observe the genetic diversity in the groups of animals generated with K-means. The first two dimensions of the PCA explained 22.21% to 58.9% of the variation found in the markers associated with mastitis, Figure 6. These results were higher than those reported in similar studies [37,38]. These authors found that 10.9% and 8.79%, respectively, of the variation was explained by the first two dimensions.

For traits RM and RSM, which possess a larger number of the loci under study, the number of eigenvalues considered to reach a minimum of 80% of variability is higher than that for the rest of the traits. However, the first two dimensions possess a high percentage of variation, evident in the graphic representation of the PCA, where the K-means groups are clearly differentiated, particularly for the RCM graph.

For RM, the genes ARHGAP10 and BDH2 (SNPs 2 and 3, respectively) contributed the most to the analysis; however, the relationship between them is negative (Figure 7). This means the genotypes for the genes in an animal are present as AA and in another as BB, or vice versa. Between these genes there was no known linkage or genetic correlation, given that they were in different chromosomes and affect unrelated traits. ARHGAP10 is associated with intramuscular fat formation [52] and BDH2 with feed intake [53].

.

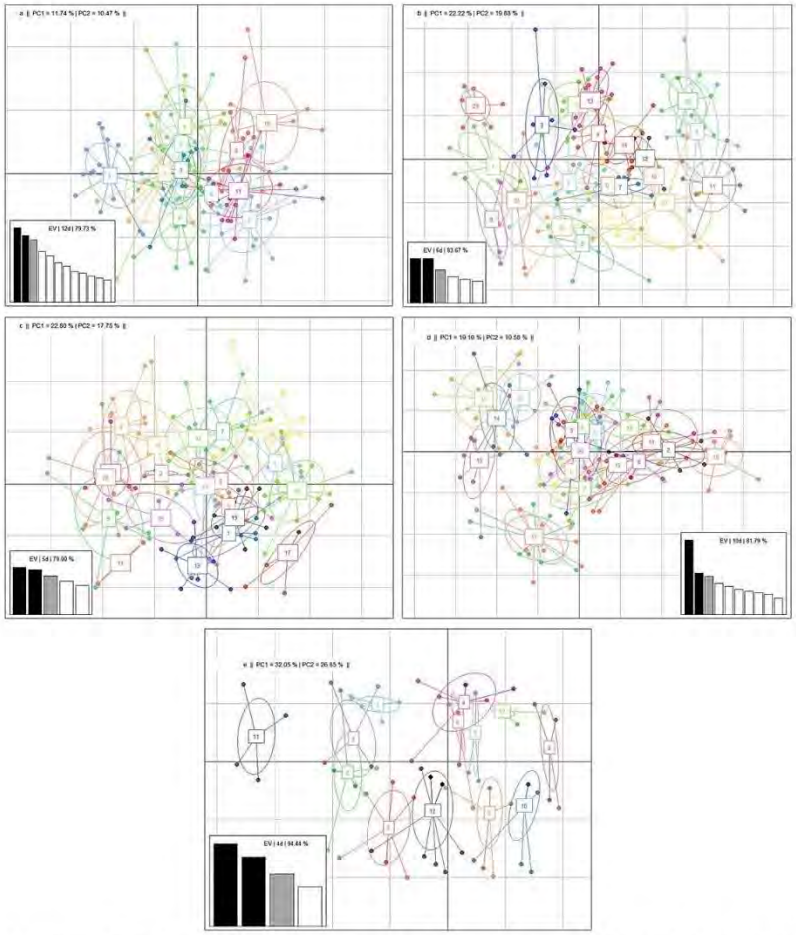

Figure 6. Graphic representation of the Principal Components Analysis: K-means' grouping of individuals showing eigenvalues (EV) for the dimensions (d) whose sum of variability is around 80 %, for resistance (a) and susceptibility (b) to mastitis, resistance to bacterial mastitis (c), subclinical mastitis (d) and clinical mastitis (e) in a Mexican Braunvieh cattle population. PC1 = percentage of variability explained by dimension one. PC2 = percentage of variability explained by dimension two.

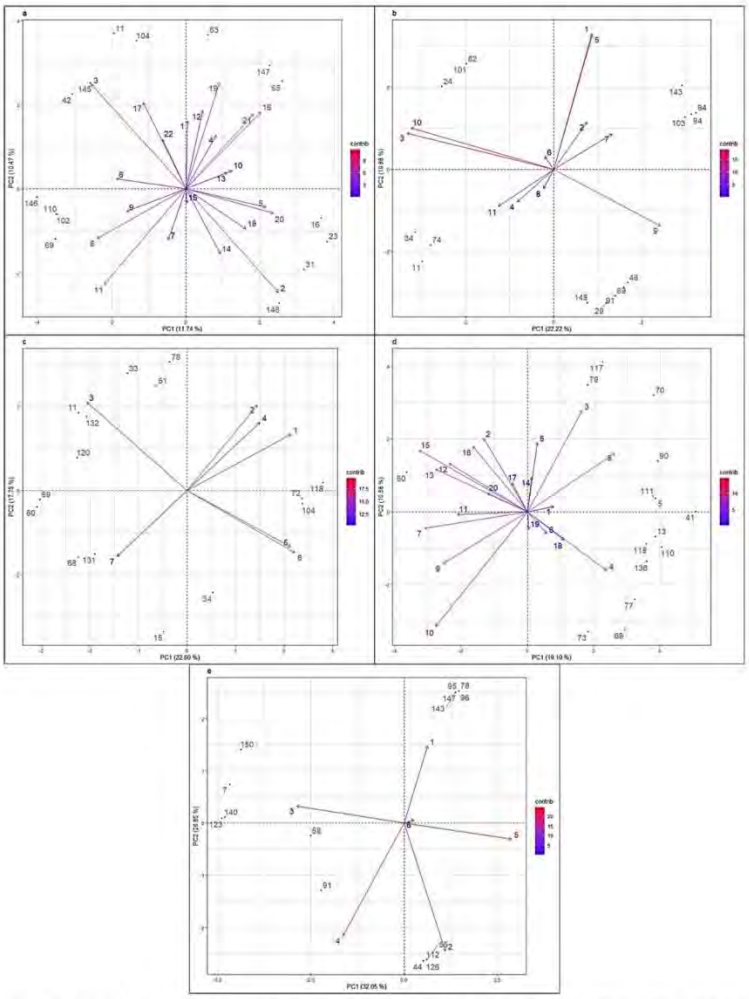

Figure 7. Biplots: marker's contribution and 10% most contributing individuals for resistance (a) and susceptibility (b) to mastitis, resistance to bacterial mastitis (c), subclinical mastitis (d) and clinical mastitis (e) in a Mexican Braunvieh cattle population. PC1 = percentage of variability explained by dimension one. PC2 = percentage of variability explained by dimension two. Note: the names of each locus by trait are indicated in Tables 1 to 5.

Oddly enough, for RM, the highly inbred animal 16 is among the top 10% of the most contributory animals (Figure 7). This can probably be explained by its position in the biplot, which means it is most associated with gene TBC1D8, which contributed little in the analysis. The same situation is true for SM, animal 143, and the marker for the gene ITPK1.

For SM, the markers for genes ANKRD33B, CTNND2, GRIA3, and SIDT1 contributed the most to the PCA, Figure 6. The pairs of genes CTNND2 with SIDT1, and ANKRD33B

with GRIA3 had a positive relationship (Figure 7) even though they are in distant chromosomes. This may be a particular feature of the genotypic profile for mastitis in this Braunvieh population.

The markers of genes IL1A and ITGB3 for RBM were in a relationship similar to that of the genes ARHGAP10 and BDH2 for RM, Figure 6. They had a negative relationship according to the biplot, Figure 7. Similarly, they are in different chromosomes, even though both directly influence the function of the immune system [19-20].

The marker BovineHD0300029925 (10) of HIVEP3, Figures 6 and 7, contributed the most to the PCA for RSM. This SNP is positively related to another two intragenic HIVEP3 markers: BovineHD0300029904 (9) and BovineHD0300029955 (11). For RCM, most of the markers were highly contributive, i.e., the SNPs for the genes CXCL8, SEL1L, and STAT4, associated directly with the immune system [47].

4. Conclusions

The population studied presented high genetic diversity for resistance and susceptibility to mastitis, even higher than that of the original Braunvieh and Brown Swiss. Although 13.3% were highly inbred animals (>50%), there were highly heterozygous groups in terms of the traits studied, a favorable indicator of the presence of diversity. Even though, the homozygous genotypes for immunologic genes (group CXCL) might be favorable for traits such as resistance to subclinical mastitis and resistance to clinical mastitis.

The K-means algorithm and the Principal Components Analysis are good statistical tools to visualize the groups of animals with different genotypic profiles. The Principal Components Analysis was useful to detect relationships and contributions of markers for each trait. The results of our research give evidence of the genetic potential of the Mexican Braunvieh population for improvement of mastitis-related traits in the tropic.

Author Contributions: Conceptualization, A.R.F. and M.Z.T.C.; methodology, M.Z.T.C.; software, P.P.R.; validation, P.P.R.; formal analysis, M.Z.T.C. and R.S.R.; investigation, A.R.F. and M.Z.T.C.; resources, A.R.F.; data curation, A.R.F.; writing—original draft preparation, M.Z.T.C. and R.S.R.; writing—review and editing, A.R.F. and P.P.R.; visualization, A.R.F.; supervision, A.R.F. and P.P.R.; project administration, A.R.F.; funding acquisition, A.R.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by Universidad Autónoma Chapingo, grant number DGIP-166701012.

Institutional Review Board Statement: Ethical review and approval were waived for this study because the producers obtained the hair samples for genotyping at their farms following their procedures.

Informed Consent Statement: Not applicable.

Data Availability Statement: Upon request to the author for correspondence.

Acknowledgments: Authors thank the Consejo Nacional de Ciencia y Tecnología, Mexico, 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.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Cobirka, M.; Tancin, V.; Slama, P. Epidemiology and Classification of Mastitis. *Animals* **2020**, 10, 2212. https://doi.org/10.3390/ani10122212
2. Suleiman, T.S.; Karimuribo, E.D.; Mdegela, R.H. Prevalence of bovine subclinical mastitis and antibiotic susceptibility patterns of major mastitis pathogens isolated in Unguja island of Zanzibar, Tanzania. *Trop. Anim. Health Prod.* **2018**, 50, 259–266. https://doi.org/10.1007/s11250-017-1424-3

3. Wang, N.; Zhou, C.; Basang, W.; Zhu, Y.; Wang, X.; Li, C.; Chen, L.; Zhou, X. Mechanisms by which mastitis affects reproduction in dairy cow: A review. *Reprod. Domest. Anim.* **2021**, 56, 1165-1175. https://doi.org/10.1111/rda.13953
4. Meuwissen, T.H.E.; Hayes, B.J.; Goddard M.E. Genomic selection: A paradigm shift in animal breeding. *Anim. Front.* **2016**, 6, 6-14. https://doi.org/10.2527/af.2016-0002
5. Haugaard, K.; Tusell, L.; Perez, P., Gianola, D.; Whist, A.C.; Heringstad, B. Prediction of clinical mastitis outcomes within and between environments using whole-genome markers. *J. Dairy Sci.* **2013**, 96, 3986–3993. https://doi.org/10.3168/jds.2012-6133
6. Brunes, L.C.; Baldi, F.; Lopes, F.B.; Narciso, M.G.; Lobo, R.B.; Espigolan, R.; Costa, M.F.O.; Magnabosco, C.U. Genomic prediction ability for feed efficiency traits using different models and pseudo-phenotypes under several validation strategies in Nelore cattle. *Animal* **2021**, 15, 100085. https://doi.org/10.1016/j.animal.2020.100085
7. Fang, L.; Sahana, G.; Ma, P.; Su, G.; Yu, Y.; Zhang, S.; Lund, M. S.; Sørensen, P. Use of biological priors enhances understanding of genetic architecture and genomic prediction of complex traits within and between dairy cattle breeds. *BMC Genom.* **2017**, 18, 604. https://doi.org/10.1186/s12864-017-4004-z
8. Meier, S.; Arends, D.; Korkuć, P.; Neumann, G.B.; Brockmann, G.A. A genome-wide association study for clinical mastitis in the dual-purpose German Black Pied cattle breed. *J. Dairy Sci.* **2020**, 103, 10289-10298. https://doi.org/10.3168/jds.2020-18209
9. Wiggans, G.R.; Cole, J.B.; Hubbard, S.M.; Sonstegard, T. S. Genomic selection in dairy cattle: the USDA experience. *Annu. Rev. Anim. Biosci.* **2017**, 5, 309-327. https://doi.org/10.1146/annurev-animal-021815-111422
10. Ali, I.; Suhail, S.M.; Shafiq, M. Heritability estimates and genetic correlations of various production and reproductive traits of different grades of dairy cattle reared under subtropical condition. *Reprod. Domest. Anim.* **2019**, 54, 1026-1033. https://doi.org/10.1111/rda.13458
11. Bhati, M.; Kadri, N. K.; Crysnanto D.; Pausch, H. Assessing genomic diversity and signatures of selection in Original Braunvieh cattle using whole-genome sequencing data. *BMC Genom.* **2020**, 21, 27. https://doi.org/10.1186/s12864-020-6446-y
12. Xia, X.; Yao, Y.; Li, C.; Zhang, F.; Qu, K.; Chen, H.; Huang, B.; Lei, C. Genetic diversity of Chinese cattle revealed by Y-SNP and Y-STR markers. *Anim. Genet.* **2018**. https://doi.org/10.1111/age.12742
13. Cai, Z.; Guldbrandtsen, B.; Lund, M. S.; Sahana, G. Prioritizing candidate genes post-GWAS using multiple sources of data for mastitis resistance in dairy cattle. *BMC Genom.* **2018**, 19, 656. https://doi.org/10.1186/s12864-018-5050-x
14. Chen, X.; Cheng, Z.; Zhang, S.; Werling, D.; Wathes, C. Combining genome wide association studies and differential gene expression data analyses identifies candidate genes affecting mastitis caused by two different pathogens in the dairy cow. *Open J. Anim. Sci.* **2015**, 5, 358-393. https://doi.org/10.4236/ojas.2015.54040
15. Kurz, J.P.; Yang, Z.; Weiss, R.B.; Wilson, D.J.; Rood, K.A.; Liu, G.E.; Wang, Z. A genome-wide association study for mastitis resistance in phenotypically well-characterized Holstein dairy cattle using a selective genotyping approach. *Immunogenetics* **2019**, 71, 35-47. https://doi.org/10.1007/s00251-018-1088-9
16. Moretti, R.; Soglia, D.; Chessa, S.; Sartore, S.; Finocchiaro, R.; Rasero, R.; Sacchi, P. Identification of SNPs associated with somatic cell score in candidate genes in Italian Holstein Friesian bulls. *Animals* **2021**, 11, 366. https://doi.org/10.3390/ani11020366
17. Meredith, B.K.; Berry, D.P.; Kearney, F.; Finlay, E.K.; Fahey, A.G.; Bradley, D.G.; Lynn, D. J. A genome-wide association study for somatic cell score using the Illumina high-density bovine beadchip identifies several novel QTL potentially related to mastitis susceptibility. *Front. Genet.* **2013**, 4, 1-10. https://doi.org/10.3389/fgene.2013.00229
18. Wang, X.; Ma, P.; Liu, J.; Zhang, Q.; Zhang, Y.; Ding, X.; Jiang, L.; Wang, Y.; Zhang, Y.; Sun, D.; Zhang, S.; Su, G.; Yu, Y. Genome-wide association study in Chinese Holstein cows reveal two candidate genes for somatic cell score as an indicator for mastitis susceptibility. *BMC Genet.* **2015**, 16, 111. https://doi.org/10.1186/s12863-015-0263-3
19. Lewandowska-Sabat, A.M.; Kirsanova, E.; Klopp, C.; Solberg, T.R.; Heringstad, B.; Østerås, O.; Boysen, P.; Olsaker, I. Transcription profiling of monocyte-derived macrophages infected in vitro with two strains of Streptococcus agalactiae reveals candidate pathways affecting subclinical mastitis in cattle. *Front. Genet.* **2019**, 10, 689. https://doi.org/10.3389/fgene.2019.00689
20. Li, L.; Chen, X.; Chen, Z. Identification of key candidate genes in dairy cow in response to Escherichia coli mastitis by bioinformatical analysis. *Front. Genet.* **2019**, 10, 1251. https://doi.org/10.3389/fgene.2019.01251
21. Song, M.; He, Y.; Zhou, H.; Zhang, Y.; Li, X.; Yu, Y. Combined analysis of DNA methylome and transcriptome reveal novel candidate genes with susceptibility to bovine Staphylococcus aureus subclinical mastitis. *Sci. Rep.* **2016**, 6, 29390. https://doi.org/10.1038/srep29390
22. Kirsanova, E.; Boysen, P.; Johansen, G.M.; Heringstad, B.; Lewandowska-Sabat, A.; Olsaker, I. Expression analysis of candidate genes for chronic subclinical mastitis in Norwegian Red cattle. *J. Dairy Sci.* **2020**, 103, 9142-9149. https://doi.org/10.3168/jds.2020-18170
23. 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 June 15, 2021.
24. 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/
25. Sherwin, W.B. Entropy and information approaches to genetic diversity and its expression: Genomic geography. *Entropy* **2010**, 12, 1765–1798. https://doi.org/10.3390/e12071765
26. Jombart, T.; Ahmed, I. Adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics* **2011**, 27, 3070-3071. https://doi.org/10.1093/bioinformatics/btr521
27. Paradis, E. Pegas: An R package for population genetics with an integrated-modular approach. *Bioinformatics* **2010**, 26, 419-420. https://doi.org/10.1093/bioinformatics/btp696

28. Kassambara, A.; Mundt, F. Factoextra: Extract and visualize the results of multivariate data analyses. R package version 1.0.7. **2020**. https://CRAN.R-project.org/package=factoextra
29. Lengyel, A.; Botta-Dukát, Z. Silhouette width using generalized mean —A flexible method for assessing clustering efficiency. *Ecol. Evol.* **2019**, 9, 13231–13243. https://doi.org/10.1002/ece3.5774
30. Ward, J.H. Hierarchical grouping to optimize an objective function. *J. Am. Stat. Assoc.* **1963**, 58, 236–244. https://doi.org/10.1080/01621459.1963.10500845
31. Zeebarree, D.Q.; Haron, H.; Abdulazeez, A.M.; Zeebaree, S.R.M. Combination of K-means clustering with genetic algorithm: A review. *Int. J. Appl. Eng. Res.* **2017**, 12, 14238-14245.
32. Mrode, R.; Ojango, J.M.K.; Okeyo, A.M.; Mwacharo, J.M. Genomic selection and use of molecular tools in breeding programs for indigenous and crossbred cattle in developing countries: Current status and future prospects. *Front. Genet.* **2019**, 9, 694. https://doi.org/10.3389/fgene.2018.00694
33. Zepeda-Batista, J.L.; Parra-Bracamonte, G.M.; Núñez-Dominguez, R.; Ramírez-Valverde, R.; Ruíz-Flores, A. Screening genetic diseases prevalence in Braunvieh cattle. *Trop. Anim. Health Prod.* **2019**, 51, 25-31. https://doi.org/10.3389/10.1007/s11250-018-1655-y
34. Trujano-Chavez, M. Z.; Valerio-Hernández, J.E.; López-Ordaz, R.; Pérez-Rodríguez, P.; Ruíz-Flores, A. Allelic and genotypic frequencies for loci associated with meat quality in Mexican Braunvieh cattle. *Trop. Anim. Health Prod.* **2021**, 53, 307. https://doi.org/10.1007/s11250-021-02757-5
35. Champely, S. pwr: Basic Functions for Power Analysis. R package version 1.3-0 **2020**. https://CRAN.R-project.org/package=pwr
36. Khalilzadeh, J.; Tasci, A.D.A. Large sample size, significance level, and the effect size: Solutions to perils of using big data for academic research. *Tour. Manag.* **2017**, 62, 89-96. https://doi.org/10.1016/j.tourman.2017.03.026
37. Upadhyay, M.; Eriksson, S.; Mikko, S.; Strandberg, E.; Stålhammar, H.; Groenen, M.A.M.; Crooijmans, P.M.A.; Anderson, G.; Johansson, A.M. Genomic relatedness and diversity of Swedish native cattle breeds. *Genet. Sel. Evol.* **2019**, 51, 56. https://doi.org/10.1186/s12711-019-0496-0
38. Zhang, W.; Gao, X.; Zhang, Y.; Zhao, Y.; Zhang, J.; Jia, Y.; Zhu, B.; Xu, L.; Zhang, L.; Gao, H.; Li, J.; Chen, Y. Genome-wide assessment of genetic diversity and population structure insights into admixture and introgression in Chinese indigenous cattle. BMC *Genet.* **2018**, 19, 114. https://doi.org/10.1186/s12863-018-0705-9
39. Agung, P.P.; Saputra, F.; Zein, M.; Wulandari, A.S.; Putra, W.; Said, S.; Jakaria, J. Genetic diversity of Indonesian cattle breeds based on microsatellite markers. *Asian-australas. J. Anim. Sci.* **2019**, 32, 467–476. https://doi.org/10.5713/ajas.18.0283
40. Abdelmanova, A.S.; Kharzinova, V.R.; Volkova, V.V.; Mishina, A.I.; Dotsev, A.V.; Sermyagin, A.A.; Boronetskaya, O.I.; Petrikeeva, L.V.; Chinarov, R.Y.; Brem, G.; Zinovieva, N.A. Genetic diversity of historical and modern populations of Russian cattle breeds revealed by microsatellite analysis. *Genes* **2020**, 11, 940. https://doi.org/10.3390/genes11080940
41. Ocampo, R.J.; Martínez, J.F.; Martínez, R. Assessment of genetic diversity and population structure of Colombian Creole cattle using microsatellites. *Trop. Anim. Health Prod.* **2021**, 53, 122. https://doi.org/10.1007/s11250-021-02563-z
42. Moscarelli, A.; Sardina, M.T.; Cassandro, M.; Ciani, E.; Pilla, F.; Senczuck, G.; Portolano, B.; Mastrangelo, S. Genome-wide assessment of diversity and differentiation between original and modern Brown cattle populations. *Anim. Genet.* **2020**, 52, 21-31. https://doi.org/10.1111/age.13019
43. Freitas, P.; Wang, Y.; Yan, P.; Oliveira, H.; Schenkel, F.; Zhang, Y.; Xu, Q.; Brito, L. Genetic diversity and signatures of selection for thermal stress in cattle and other two Bos Species adapted to divergent climatic conditions. *Front. Genet.* **2021**, 12, 604823. https://doi.org/10.3389/fgene.2021.604823
44. Saravanan, K.A.; Panigrahi, M.; Kumar, H.; Parida, S.; Bhushan, B.; Gaur, G.K.; Kumar, P.; Dutt, T.; Mishra, B.P.; Singh, R.K. Genome-wide assessment of genetic diversity, linkage disequilibrium and haplotype block structure in Tharparkar cattle breed of India. *Anim. Biotechnol.* **2020**, 1–15. https://doi.org/10.1080/10495398.2020.1796696
45. Strucken, E.M.; Gebrehiwot, N.Z.; Swaminathan, M.; Joshi, S.; Al Kalaldeh, M.; Gibson, J.P. Genetic diversity and effective population sizes of thirteen Indian cattle breeds. *Genet. Sel. Evo.* **2021**, 53, 47. https://doi.org/10.1186/s12711-021-00640-3
46. King, F.J.M.; Banga, C.B.; Visser, C. Genetic diversity and population structure of three native cattle populations in Mozambique. *Trop. Anim. Health Prod.* **2021**, 53, 117. https://doi.org/10.1007/s11250-021-02562-0
47. Robbertse, L.; Richards, S.A.; Maritz-Olivier, C. Bovine immune factors underlying tick resistance: Integration and future directions. *Front. Cell. Infect. Microbiol.* **2017**, 7, 522. https://doi.org/10.3389/fcimb.2017.00522
48. Kiser, J.N.; Wang, Z.; Zanella, R.; Scraggs, E.; Neupane, M.; Cantrell, B.; Van Tassell, C.P.; White, S.N.; Taylor, J.F.; Neibergs, H. L. Functional variants surrounding Endothelin 2 are associated with Mycobacterium avium subspecies paratuberculosis infection. *Front. Vet. Sci.* **2021**, 8, 280. https://doi.org/10.3389/fvets.2021.625323
49. Signer-Hasler, H.; Burren, A.; Neuditschko, M.; Frischknecht, M.; Garrick, D.; Stricker, C.; Gredler, B.; Bapst, B.; Flury, C. Population structure and genomic inbreeding in nine Swiss dairy cattle populations. *Genet. Sel. Evo.* **2017**, 49, 83. https://doi.org/10.1186/s12711-017-0358-6
50. Balloux, F.; Amos, W.; Coulson, T. Does heterozygosity estimate inbreeding in real populations? *Mol. Ecol.* **2004**, 13, 3021–3031. https://doi.org/10.1111/j.1365-294x.2004.02318.x
51. Tsairidou, S.; Allen, A.R.; Pong-Wong, R.; McBride, S.H.; Wright, D.M.; Matika, O.; Pooley, C.M.; McDowell, S.W.J.; Glass, E.J.; Skuce, R.A.; Bishop, S.C.; Woolliams, J.A. An analysis of effects of heterozygosity in dairy cattle for bovine tuberculosis resistance. *Anim. Genet.* **2018**, 49, 103–109. https://doi.org/10.1111/age.12637

52. Ueda, S.; Hosoda, M.; Yoshino, K.; Yamanoue, M.; Shirai, Y. Gene expression analysis provides new insights into the mechanism of intramuscular fat formation in Japanese Black Cattle. *Genes* **2021**, 12, 1107. https://doi.org/10.3390/genes12081107
53. Salleh, M.S.; Mazzoni, G.; Nielsen, M.O.; Løvendahl, P.; Kadarmideen, H.N. Identification of expression QTLs targeting candidate genes for residual feed intake in dairy cattle using systems genomics. *J. Genet. Genome. Res.* **2018**, 5, 38. https://doi.org/10.23937/2378-3648/1410035

CAPÍTULO 6. CONCLUSIONES GENERALES

Los capítulos 3 a 5 muestran evidencia de la alta diversidad genética en bovinos Suizo Europeo para características complejas, sinónimo de su versatilidad genética sobresaliente en comparación con sus razas emparentadas Braunvieh original y Pardo Suizo; debido a que esta población se ha desarrollado por más de un siglo bajo condiciones tropicales. La diversidad genética se midió mediante la heterocigosidad observada, coeficiente de consanguinidad, frecuencia de alelos favorables y estadística multivariada.

En el capítulo 3 se encontraron frecuencias mayores a las de otras razas para genotipos asociados con mayor suavidad, mayor marmoleo, menor cantidad de ácidos grasos saturados, y menor fuerza de corte (CAPN1 y DGAT2). Existen similitudes con las frecuencias de la mayoría de las razas Europeas para los genes CAPN3 y TG; mientras que, para los genes DGAT1 y ANK1 las frecuencias de los genotipos deseados son inferiores. Un marcador de DGAT1 y el marcador de MADH3 resultaron monomórficos.

En el estudio para características reproductivas (capítulo 4), los valores de Ho cercanos a 0.5, sugieren un incremento en su versatilidad debido al ambiente tropical, dado que el valor difiere al de los ancestros directos de la raza y coincide con los de razas adaptadas al trópico. Además, el gen DSC2 fue el más relevante en este estudio ya que explicó variabilidad en los grupos para todas las características reproductivas estudiadas.

Por último, en el capítulo 5, se encontró que, para resistencia y susceptibilidad a mastitis, los perfiles genotípicos mostraron alta diversidad genética entre grupos, donde resistencia a mastitis clínica tuvo la menor presencia de genotipos heterocigóticos. Aunque el porcentaje de animales altamente consanguíneos (>50%) es de hasta 13.3%, existen grupos altamente heterocigóticos para las características estudiadas, indicador favorable de la presencia de diversidad.

En general, el uso de estadística multivariada fue clave para encontrar a los grupos de animales con perfiles genotípicos similares y el número de animales heterocigóticos por característica, lo anterior fue útil para determinar el grado de diversidad genética de la población para las diferentes características estudiadas. Los resultados de estas investigaciones evidencian el potencial genético de la población Suizo Europeo mexicana, para mejorar características relacionadas con calidad de la carne, reproducción y resistencia a mastitis.