UNIVERSIDAD AUTÓNOMA CHAPINGO

DEPARTAMENTO DE ENSEÑANZA, INVESTIGACIÓN
Y SERVICIO EN ZOOTECNIA
POSGRADO EN PRODUCCIÓN ANIMAL

POLIMORFISMOS EN GENES CANDIDATOS ASOCIADOS A CARACTERES ECONÓMICAMENTE IMPORTANTES EN BOVINOS SUIZO EUROPEO

TESIS

Que como requisito parcial
para obtener el grado de:

DOCTOR EN CIENCIAS EN INNOVACIÓN GANADERA

Presenta:
JOSÉ LUIS ZEPEDA BATISTA

Bajo la supervisión de: RAFAEL NÚÑEZ DOMÍNGUEZ, Ph.D.
GASPAR MANUEL PARRA BRACAMONTE, Dr.

Junio de 2018

Chapingo, Estado de México

POLIMORFISMOS EN GENES CANDIDATOS ASOCIADOS A CARACTERES ECONÓMICAMENTE IMPORTANTES EN BOVINOS SUIZO EUROPEO

Tesis realizada por **JOSÉ LUIS ZEPEDA BATISTA** bajo la supervisión del Comité Asesor indicado, aprobada por el mismo y aceptada como requisito parcial para obtener el grado de:

DOCTOR EN CIENCIAS EN INNOVACIÓN GANADERA

DIRECTOR:

Ph.D. RAFAEL NÚÑEZ DOMÍNGUEZ

CO-DIRECTOR:

Dr. GASPAR MANUEL PARRA BRACAMONTE

ASESOR:

Ph.D. RODOLFO RAMÍREZ VALVERDE

ASESOR:

Ph.D. AGUSTÍN RUÍZ FLORES

LECTOR EXTERNO:

Ph.D. FELIPE ALONSO RODRÍGUEZ ALMEIDA

CONTENIDO

LISTA DE CUADROS

LISTA DE CUADROS (Continuación)

LISTA DE FIGURAS

LISTA DE FIGURAS (Continuación)

DEDICATORIAS

Al creador, porque me permitió salir avante de las turbulentas aguas que me tocó navegar. Y es que mientras más aprendo con la ciencia, más me convenzo de la existencia del creador.

A nuestra señora de San Juan de los Lagos, por ayudarme a tomar rumbo durante la fase más complicada del Doctorado, así como demostrarme que la fe mueve montañas, retira al enemigo de nuestro lado y allana el camino para el paso del viajero.

A mis padres, les ofrezco este escrito como muestra de gratitud por todo el apoyo y confianza que siempre me han brindado. Espero no haber defraudado sus expectativas. ¡Los quiero muchísimo!

A mi esposa Isabel, porque es la persona que mejor ha sabido entenderme en el desorden al que llamo vida. Nunca imaginé que esa vida me permitiera recorrer el hilo del destino para lograr conocerte, mi alma gemela. ¡Te amo, hermosa!

A mis hijos José Luis y Sofía Isabel, porque desde que llegaron a mí, han sido motores de vida. Porque muy a su manera, siempre me han dado muchos ánimos para continuar. Y sobre todo porque han hecho de su padre, una mucho mejor persona.

A mis familiares, por todas las buenas vibras, consejos, y porras que tuvieron a bien darme en cada reunión. Sobre todo, a mi hermana Olimpia, mi confidente de la niñez y mi arreadora de la adultez. ¡Te quiero un buen morra!

A los amigos que he obtenido a lo largo de mi vida, porque me mostraron la multitud de caras que puede adoptar la realidad. Porque en su momento compartieron tristezas, enojos, alegrías y… Una que otra fiestecita. ¡Los aprecio y nunca los olvidaré "compares"!

A todos mis profesores, porque tuvieron a bien tomarse la molestia de moldear esa arcilla que llegó a sus manos. Porque el compromiso y la vocación por su trabajo los hizo comprender que aquel niño, joven, o adulto, alegre, inquieto e hiperactivo, no era mal estudiante, simplemente tenía un modo muy peculiar para desempeñar sus actividades. A todos ustedes los aprecio, los recuerdo, y sobre todo ¡los admiro! ¡Muchas gracias por todas sus enseñanzas!

AGRADECIMIENTOS

A la Universidad Autónoma Chapingo, por la oportunidad de hacer mis estudios de doctorado.

Al Posgrado en Producción Animal por darme el privilegio de seguir formándome académicamente.

Al Consejo Nacional de Ciencia y Tecnología por el financiamiento dado durante mis estudios de nivel doctorado.

A la Asociación Mexicana de Criadores de Ganado Suizo de Registro por la oportunidad de trabajar con sus socios.

Al Dr. Rafael Núñez Domínguez por sus valiosos consejos y echarse a la espalda la responsabilidad que significó el continuar con mi programa doctoral. Así como por su amistad y palabras de aliento.

Al Dr. Gaspar Manuel Parra Bracamonte por la paciencia, dirección, supervisión y enseñanza en todo momento durante la estancia que realicé en Reynosa. Así mismo por los valiosos consejos y la amistad brindados.

Al Dr. Rodolfo Ramírez Valverde, por sus excelentes consejos y observaciones durante el doctorado; así como por su amistad y las molestias que siempre se ha tomado para moldear una mejor persona.

Al Dr. Agustín Ruíz Flores por su amistad, paciencia y enseñanza durante todo este proyecto que significó el Doctorado. Así como por su apoyo a mis ideas repentinas durante el doctorado.

Al cuerpo de académicos e investigadores que forman parte del PPA por la experiencia transmitida para formarnos como profesionistas de alta calidad y enfrentar la realidad del sector agropecuario.

A los “Compitas” (Saavedra, Orlando, Citlalli, Vite, Julio, Elba) y al team Reynosa (Dra. Ana, Jahuey, Paz, Xóchitl, Williams, Diana, Oliva, Paula, Hugo, Erick, Amado,
George) por la gran amistad que me brindaron, así como el apoyo tan importante que recibí de ellos en diversas ocasiones a lo largo del doctorado.
A todos los que directa o indirectamente colaboraron para que esta investigación llegara a buen término.

DATOS BIOGRÁFICOS

Datos personales

Nombre: José Luis Zepeda Batista

Fecha de nacimiento: 25 de marzo de 1986

Lugar de nacimiento: Colima, Colima

CURP: ZEBL860325HCMPTS06

No. Cartilla Militar: 7894232

Profesión: Maestro en Ciencias en Innovación Ganadera

Cédula de Maestría: 11050410

Desarrollo académico

2014-2018	Doctor en Ciencias en Innovación Ganadera. Posgrado en Producción Animal. Departamento de Zootecnia. Universidad Autónoma Chapingo. Chapingo, México.
2011-2012	Maestro en Ciencias en Innovación Ganadera. Posgrado en Producción Animal. Departamento de Zootecnia. Universidad Autónoma Chapingo. Chapingo, México.
2004-2009	Médico Veterinario Zootecnista. Facultad de Medicina Veterinaria y Zootecnia. Universidad de Colima. Tecomán, Colima.

RESUMEN

En bovinos, la información de los rasgos económicamente relevantes (RER) es tradicionalmente incluida en los criterios de selección de los programas de cría de ganado de carne, debido a la importancia económica para ganaderos e industria. Sin embargo, la identificación de variantes genéticas causales que afecten directamente los fenotipos de los RER es una tarea difícil. El uso de miles de marcadores SNP en estudios de asociación de genoma completo (GWAS) ha permitido la identificación y confirmación de muchos QTL para caracteres de crecimiento en ganado bovino, que a su vez explican la variación fenotípica. El presente documento tuvo como objetivo la estimación de las frecuencias genotípica y génica de polimorfismos en genes candidatos asociados con RER, así como estimar el efecto de dichos genes sobre rasgos de crecimiento en bovinos Pardo Suizo Europeo. El análisis de 28 enfermedades genéticas usualmente presentes en ganado productor de carne indicó que la población mexicana de ganado Pardo Suizo Europeo no posee marcadores previamente asociados con 15 de dichas enfermedades. Por otro lado, se determinó la presencia de animales heterocigotos y homocigotos para marcadores asociados con 13 de las 28 enfermedades genéticas estudiadas. Ninguno de los animales homocigotos mostró signos clínicos o subclínicos de las enfermedades estudiadas. Mientras tanto, los resultados del GWAS determinaron una región asociada con peso al nacimiento que no ha sido previamente reportada (BWT_rs133262280_22_60.7). Además, se identificaron dos regiones previamente asociadas con peso al destete (WW_rs43668789_11_21.3 y WW_rs136155567_27_27.0) con un efecto promedio de sustitución del alelo de 4.08 %. Los resultados obtenidos en la presente investigación indican que la “salud genética” de la población mexicana de Pardo Suizo Europeo en general es buena. Sin embargo, la presencia de animales heterocigotos para algunas de las enfermedades genéticas estudiadas hace necesario un continuo seguimiento de los animales, con el fin de evitar la posible presencia de las formas clínicas de dichas enfermedades. Por otro lado, la identificación de regiones asociadas con características de crecimiento en la población estudiada mostró la posibilidad de realizar un mejoramiento genético de la raza Pardo Suizo Europeo, utilizando tecnologías de genotipado como herramientas auxiliares de las evaluaciones genéticas que actualmente se llevan a cabo en México.

Palabras clave: enfermedades genéticas, crecimiento, peso vivo, QTL

ABSTRACT

In cattle, information of economically relevant traits (ERT) is traditionally included in the selection criteria of beef cattle breeding programs, due to their economic importance for livestock and industry. However, the identification of causal genetic variants that directly affect ERT phenotypes is a difficult task. The use of thousands of SNP markers in genome-wide association studies (GWAS) has allowed the identification and confirmation of many QTL for growth traits in cattle, which in turn explain the phenotypic variation. The objective of this research was to estimate the genotypic and gene frequencies of polymorphisms in candidate genes associated with ERT, as well as to estimate the effect of these genes on growth traits in Braunvieh cattle. The analysis of 28 genetic diseases usually present in beef cattle, indicated that the Mexican population of Braunvieh cattle does not have markers previously associated with 15 of these diseases. On the other hand, the presence of heterozygous and homozygous animals was determined for markers associated with 13 of the 28 genetic diseases studied. None of the homozygous animals showed clinical or subclinical signs of the diseases studied. Meanwhile, the results of the GWAS determined a single region associated with birth weight that has not been previously reported (BWT_rs133262280_22_60.7). Besides, two regions previously associated with weaning weight were identified (WW_rs43668789_11_21.3 and WW_rs136155567_27_27.0) with an average allele substitution effect of 4.08%. The results obtained in this research indicate that the "genetic health" of the Mexican population of Braunvieh cattle is good. However, the presence of heterozygous animals for some of the genetic diseases studied makes necessary the continuous monitoring of the animals in order to avoid the possible presence of the clinical forms of these diseases. On the other hand, the identification of regions associated with growth traits in the population studied, showed the possibility of genetic improvement of the Braunvieh breed, using genotyping technologies as auxiliary tools of the genetic evaluations that are currently carried out in Mexico.

Key words: genetic diseases, growth, live weight, QTL

1. INTRODUCCIÓN GENERAL

En bovinos, las características de crecimiento son tradicionalmente incluidas en los criterios de selección de los programas de cría de bovinos para carne, debido a su asociación con la producción de carne; por tanto, son de gran importancia económica para los ganaderos y la industria (Barwick & Henzell, 2005). Dichas características económicamente relevantes (CER) son consideradas genéticamente complejas, por lo cual, se han utilizado diferentes enfoques para identificar la variación genética relacionada con las diferencias fenotípicas. Sin embargo, la identificación de variantes genéticas causales que afecten directamente los fenotipos de las CER es una tarea difícil (Zhu & Zhao, 2007). Aunque la asociación de genoma completo se ha convertido en la estrategia más frecuentemente aplicada para identificar la variación genética que influyen las CER, el enfoque de genes candidatos también ha sido ampliamente utilizado para identificar dicha variación.

La estrategia de gen candidato consiste en buscar y utilizar genes que se espera estén asociados con la expresión de una característica y define si la variación genética presente en las poblaciones se asocia con la diversidad fenotípica, siendo además altamente eficiente debido a la identificación de genes específicos a lo largo del genoma (Zhu *& Zhao*, 2007). En un estudio de asociación, dos de los pasos críticos utilizados en el enfoque de genes candidatos son seleccionar un gen candidato adecuado e identificar las variantes genéticas o polimorfismos más útiles para la prueba.

Comúnmente, la función fisiológica, la clonación posicional y la comparación de enfoques genómicos se han utilizado para seleccionar los genes candidatos (Lindholm-Perry *et al*., 2012; Morsci *et al*., 2006; Womack, 2005); sin embargo, el análisis de redes de interacción o redes neuronales también pueden ser una

excelente alternativa para la selección de genes candidatos por los equipos de expertos en bovinos. Por tanto, en la ciencia animal se ha comenzado a utilizar herramientas bioinformáticas para modelar y generar redes de interacción que representan la genética de los rasgos complejos arquitectónicos en bovinos, tales como marmoleo, edad de la pubertad y características reproductivas (Fortes *et al*., 2011; Hulsegge *et al*., 2013; Lim *et al*., 2011). En este sentido, Snelling *et al.* (2010) reportaron importantes asociaciones genómicas para pesos al nacimiento, al destete y al año en la raza brasileña Canchim, abriendo así una oportunidad para la realización de mejoramiento genético con enfoque genómico.

En México, la crianza de bovinos para producción de carne se realiza con razas especializadas y para doble propósito, siendo Suizo Europeo la más extendida a nivel nacional (CONARGEN, 2016). Esta raza ha mostrado un buen desempeño productivo de acuerdo con investigaciones previamente realizadas para evaluar algunos parámetros genéticos productivos (Ruíz-Flores *et al*., 2006; Saavedra-Jiménez *et al*., 2013). Sin embargo, las condiciones de crianza de la raza y los intervalos generacionales actuales han restringido la velocidad de avance genético obtenido mediante las evaluaciones genéticas tradicionales.

Por otro lado, en México, las evaluaciones genéticas auxiliadas por marcadores moleculares, tales como los SNP, es un área de oportunidad ya que su desarrollo es incipiente. Por lo anterior, la implementación de los resultados de este estudio permitiría un incremento en el avance genético de la raza, innovando además en el mejoramiento genético de la ganadería bovina productora de carne.

1.1. Hipótesis

Los sesgos en las metodologías de asociación genética podrán ser reducidos al ser implementadas las estrategias establecidas durante la revisión de literatura.

El análisis de enfermedades genéticas en la población de Suizo Europeo mexicano mostrará pocos animales homocigotos para la mayoría de los loci relacionados con enfermedades.

El análisis de genoma completo permitirá identificar genes candidatos asociados con características de crecimiento, que puedan utilizarse en programas de selección y mejoramiento genético.

1.2. Objetivos

Identificar las fuentes de sesgo en los análisis de asociación genética reportados en la literatura, para establecer estrategias de reducción de dichas fuentes de sesgo en el presente estudio.

Estimar la frecuencia genotípica y alélica de polimorfismos en genes candidatos asociados con enfermedades genéticas en bovinos Pardo Suizo Europeo.

Estimar el efecto de polimorfismos en genes candidatos en características de crecimiento en bovinos Pardo Suizo Europeo.

1.3. Estructura de la tesis

En el primer capítulo de esta tesis se aborda de manera general el estado del conocimiento actual sobre tecnologías genómicas, su uso y aplicación. Así como las oportunidades y aplicabilidad a nivel nacional que dichas tecnologías tienen para el mejoramiento genético de la raza Pardo Suizo Europeo.

En el segundo capítulo se presenta una revisión de literatura sobre la clasificación y el efecto potencial de los sesgos en la asociación de genotipos y características productivas en el ganado, así como posibles estrategias para reducir la incidencia de dichos sesgos.

El tercer capítulo contiene un metaanálisis de la influencia de los genes de κ-caseína (*CSN3*) y β-lactoglobulina (*LGB*) en estudios de asociación gen-característica productiva, así como la comparación entre dichas asociaciones en cuatro razas bovinas.

En el cuarto capítulo se muestra una revisión de literatura sobre genes candidatos para ser utilizados en estudios de asociación genética destinados a realizar

mejoramiento genético de características de crecimiento y calidad de la carne en razas de bovinos para carne.

En el penúltimo y quinto capítulo, se presenta un barrido de marcadores asociados con enfermedades genéticas para determinar la “salud genética” de la población de Suizo Europeo mexicana. Además, se identifican particularidades del comportamiento de dichos marcadores que permitirán utilizarlos en programas de selección y mejoramiento genético.

Por último, en el sexto capítulo se describen los resultados obtenidos de un análisis de genoma completo, orientado a localizar regiones del genoma asociadas con características de crecimiento en la población de Suizo Europeo mexicana. Además, se identificaron genes candidatos asociados con características de importancia económica, que podrán aplicarse en programas de selección asistida por marcadores en bovinos Suizo Europeo mexicanos.

1.4. Literatura citada

Barwick, S. A., & Henzell, A. L. (2005). Development successes and issues for the future in deriving and applying selection indexes for beef breeding. *Australian Journal of Experimental Agriculture,* 45, 923-933.

CONARGEN (Consejo Nacional de los Recursos Genéticos Pecuarios). (2016). Disponible en (http://www.conargen.mx/index.php/asociaciones/bovinos-leche). Consultada el 28 de enero de 2016.

Fortes, M. R., Reverter, A., Nagaraj, S. H., Zhang, Y., Jonsson, N. N., Barris, W., Lehnert, S., Boe-Hansen, G. B., & Hawken, R. J. (2011). A single nucleotide polymorphism-derived regulatory gene network underlying puberty in 2 tropical breeds of beef cattle. *Journal of Animal Science*, 89, 1669-1683.

Hulsegge, I., Woelders, H., Smits, M., Schokker, D., Jiang, L., & Sorensen, P. (2013). Prioritization of candidate genes for cattle reproductive traits, based on protein-protein interactions, gene expression, and text-mining. *Physiological Genomics*, 45, 400-406.

Lim, D. N., Kim, K., Park, H. S., Lee, S. H., Cho, Y. M., Oh, S. J., Kim, T. H., & Kim, H. (2011). Identification of candidate genes related to bovine marbling using protein-protein interaction networks. *International Journal of Biological Science*, 7, 992-1002.

Lindholm-Perry, A. K., Kuehn, L. A., Smith, T. P., Ferrell, C. L., Jenkins, T. G., Freetly, H. C., & Snelling, W. M. (2012). A region on BTA14 that includes the Positional candidate genes LYPLA1, XKR4 and TMEM68 is associated

with feed intake and growth phenotypes in cattle. *Animal Genetics*, 43, 216-219.

Morsci, N. S., Schnabel, R. D., & Taylor, J. F. (2006). Association analysis of adiponectin and somatostatin polymorphisms on BTA1 with growth and carcass traits in Angus cattle. *Animal Genetics* 37, 554-562.

Ruíz-Flores, A., Núñez-Domínguez, R., Ramírez-Valverde, R., Domínguez-Viveros, J., Mendoza-Domínguez, M., & Martínez-Cuevas E. (2006). Niveles y efectos de la consanguinidad en variables de crecimiento y reproductivas en bovinos Tropicarne y Suizo Europeo. *Agrociencia*, 40, 289-301.

Saavedra-Jiménez, L. A., Ramírez-Valverde, R., Núñez-Domínguez, R., García-Muñiz, J. G., López-Villalobos, N., & Ruíz-Flores, A. (2013). Genotype by climate interaction in the genetic evaluation for growing traits of Braunvieh cattle in Mexico. *Tropical Animal Health and Production,* 45(7), 1489-1494.

Snelling, W. M., Allan, M. F., Keele, J. W., Kuehn, L. A., McDaneld, T., Smith, T. P., Sonstegard, T. S., Thallman, R. M., & Bennett, G. L. (2010). Genome-wide association study of growth in crossbred beef cattle. *Journal of Animal Science*, 88, 837-848.

Womack, J. E. (2005). Advances in livestock genomics: opening the barn door. *Genome Research*, 15, 1699-1705.

Zhu, M., & Zhao, S. (2007). Candidate gene identification approach: progress and challenges. *International Journal of Biological Science*, 3, 420-427.

2. SOURCES OF BIAS IN GENETIC ASSOCIATION STUDIES OF CATTLE: A REVIEW

José Luis Zepeda-Batista[1], María Isabel Carrillo-Díaz[2], Luis Antonio Saavedra-Jiménez[1]

[1]Departamento de Zootecnia, Posgrado en Producción Animal, Universidad Autónoma Chapingo, Km 38.5 Carretera México-Texcoco Chapingo, México 56230.

[2]Facultad de Medicina Veterinaria y Zootecnia, Universidad de Colima, km. 40, Autopista Colima-Manzanillo, Colima, 28100.

Published in *Revista Colombiana de Ciencias Pecuarias. In Press.*

Sources of bias in genetic association studies of cattle: A review

Fuentes de sesgo en estudios de asociación genética en ganado bovino: Revisión de literatura

Fontes de viés nos estudos de associação genética em bovinos: Revisão da literatura

José L Zepeda Batista[1], MVZ, MSc; María I Carrillo Díaz[2*], IAZ, PhD; Luis A Saavedra Jiménez[1], IAZ, MSc.

[1]*Posgrado en Producción Animal, Universidad Autónoma Chapingo, Chapingo, Estado de México, México.*

[2]*Facultad de Medicina Veterinaria y Zootecnia, Universidad de Colima, Tecomán, Colima, México.*

Abstract

Background: In cattle breeding programs, genetic association studies have been increasingly used. However, inconsistent results, such as positive, negative, or absence of association, across studies restrains reproducibility and proper implementation, propitiating the appearance of bias. **Objective:** to identify and classify potential sources of bias and determine possible strategies to avoid it in the genetic association studies in cattle. **Source of bias in genetic association studies:** Genetic and genomic sources of bias include effects associated with the gene loci governing expression. Sampling-related and statistical biases are related with factors such as stratification and database size. **Strategies to correct bias in genetic association studies:** Correction strategies differ in nature. Genetic and genomic strategies are based on determining the appropriate approach to obtain and report the genetic information. Sampling-related and statistical strategies are based on grouping individuals with certain traits that lead to a reduction in heterogeneity. **Conclusion:** It is necessary to consider the methodology used in previous studies to establish a hierarchy of sources of bias and facilitate decisions on the use of tools to reduce inconsistencies in the results of future studies.

Keywords: *association estimates, genetic bias, genetic improvement, sampling-related bias, statistical bias.*

Resumen

Antecedentes: En los programas de mejoramiento genético, los estudios de asociación genética son cada vez más usados. Sin embargo, resultados inconsistentes, como positivos, negativos o la ausencia de asociación, a través de los estudios restringen la reproducibilidad y su aplicación adecuada, propiciando la aparición de sesgos. **Objetivo:** identificar y clasificar fuentes potenciales de sesgo y determinar posibles estrategias para evitarlo en los estudios de asociación genética en ganado. **Fuentes de sesgo en estudios de asociación genética:** Las fuentes genéticas y genómicas de sesgo incluyen los efectos asociados con la expresión que gobierna los loci. Los sesgos estadísticos y de muestreo están relacionados con factores como la estratificación y el tamaño de la base de datos. **Estrategias para corregir sesgos en estudios de asociación genética:** Las estrategias de

corrección difieren en naturaleza. Las estrategias genéticas y genómicas se basan en determinar el enfoque apropiado para obtener la información genética. Las estrategias estadísticas y relacionadas con el muestreo se basan en la agrupación de individuos con ciertos rasgos que conducen a una reducción de la heterogeneidad. **Conclusión.** Es necesario considerar las metodologías utilizadas en estudios previos, para establecer una jerarquía de las fuentes de sesgo y facilitar las decisiones en el uso de herramientas para reducir inconsistencias en los resultados futuros.

Palabras clave: *estimados de asociación, mejoramiento genético, sesgo de muestreo, sesgo estadístico, sesgo genético.*

Resumo

Antecedentes: Nos programas de criação de bovinos, os estudos de associação genética têm sido cada vez mais utilizados. No entanto, resultados inconsistentes, como positivos, negativos ou ausência de associação entre os estudos, restringem a reprodutibilidade e sua adequada implementação, propiciando o aparecimento de viés. **Objetivo:** identificar e classificar potenciais fontes de viés e determinar estratégias possíveis para evitá-lo nos estudos de associação genética em bovinos. **Fonte de viés em estudos de associação genética:** Fontes genéticas e genômicas do viés incluem os efeitos associados aos genes que relacionam a expressão. Os vícios estatísticos e de amostragem estão relacionados a fatores como a estratificação e o tamanho do banco de dados. **Estratégias para corrigir os viéses nos estudos de associação genética:** As estratégias de correção diferem na natureza. As estratégias genéticas e genômicas são baseadas na determinação da abordagem apropriada para obter e relatar a informação genética. As estratégias estatísticas e de amostragem baseiam-se no agrupamento de indivíduos com certos traços que levam a uma redução na heterogeneidade. **Conclusão:** É necessário considerar a metodologia utilizada em estudos anteriores para estabelecer uma hierarquia de fontes de viés e facilitar decisões sobre o uso de ferramentas para reduzir inconsistências nos resultados de estudos futuros.

Palavras chave: *estimativas de associação, melhoria genética, viés de amostragem, viés estatístico, viés genético.*

Introduction

Genetics association studies (GAS) aim to detect associations between one or more genetic polymorphism and a quantitative or discrete trait, by testing for a correlation between a specific trait and a genetic variation (Lewis and Knight, 2012). The number of genetic association studies have increased, and their assessment has become a powerful approach to identify common and rare variants underlying complex diseases (Wu *et al*., 2012), discovering causative mutations (Schwarzenbacher *et al.*, 2016), or identification of QTLs (Jahuey-Martínez *et al*., 2016) on a population. Nevertheless, inconsistencies in GAS due to the combination of factors contribute to the spurious or not consistently results (Table 1). These findings suggest that many original results could be false-positive (type I errors), or that small genetic effects were undetectable (false-negative, type II errors) (Lee, 2015). This lack of reproducibility tends to produce genetic associations without value for genetic improvement.

Table 1. Results of genetic association studies between CSN3 gene with milk yield in dairy cattle.

	Author		
	Gustavsson *et al*. (2014)	**Duifhuis-Rivera *et al*. (2014)**	**Deb *et al*. (2014)**
Sampled animals	400	202	200
Reported effect	Positive	Absence	Positive
Best genotype*	AA	N/D	AB

*Best genotype: genotype reported with the best performance for milk yield; N/D: gen-trait association absent.

Additionally, Ioannidis (2005) defined bias as the combination of design, data, analysis, and presentation factors that produce research findings otherwise should not be produced. However, most of the reviews of bias in GAS have been focused to analyze mostly the genetic factors or address some other factors as part of the genetic issues. Based on these, it is necessary to classify bias in GAS according to its nature to better understanding and to reduce the possible spurious results. Therefore, the objective of the current study was

to identify and classify potential sources of bias and determine possible strategies to avoid it in the genetic association studies in cattle.

Sources of bias in genetic association studies

Different approaches, based on related or non-related individuals, have been used to carry out the GAS (Table 2). The literature reports that some widely cited associations cannot be replicated due to the inaccuracies in the approaches used to determined them (Sagoo *et al*., 2009). In this sense, inconsistencies in GAS could be attributable to factors such as genetic, genomic, sampling-related, or statistical, which influence production traits, and contribute to the risk of false-positive results (Pärna *et al.*, 2012).

Table 2. Former and current approaches used in genetic association studies

	Approach	Advantages	Disadvantages
Family Based	TDT[1] FB-GWAS[2]	Quality control; robustness to population stratification; ability to perform genotyping quality control	Less power than pop-based GWAS; computationally demanding; not practical for late-onset diseases
Population based	Candidate polymorphism & Candidate gene	Determine if a given SNP or set of SNPs influences the trait directly; involve multiple SNPs within a single gene; capture information of the underlying genetic variability	SNPs may not serve as the true trait-causing variants; multiple SNPs measurements are needed to know a precise location on the genome
	Fine mapping	Set out to identify with a high level of precision the location of a trait-causing variant; determine the position on the genome of the causative mutation	Data preprocessing extensive and computational burden greater; specific software requirements; need for candidate gene studies to validate findings from GWAS
	Genome-wide	Identify associations between SNPs and a trait; involves the characterization of larger number of SNPs	

[1]TDT: transmission disequilibrium test; [2]FB-GWAS: family-based genome-wide association study (Benyamin *et al*., 2009; Foulkes, 2009).

Genetic factors

Considered breed(s) in the study could be a source of bias due to intra- and inter-racial bovine genetic population diversity (Figure 1). Besides, presence of crossbred populations, confers changes in the behavior of offspring, relative to that of the parents. Modifications can be evaluated by direct, maternal effects and heterosis of breeds and their crosses, with enough precision to predict the expected behavior of several breeding alternatives and mating systems (Dickerson, 1993). On this regard, Trail *et al*. (1984) reported direct and maternal effects on economical production traits in crossbred Boran cattle showing differences due to paternal or maternal breed.

Contemporary group (CG) is other genetic factor of bias, it affects results due to the influence of the interaction between genotype and environment (Ramírez-Valverde *et al*., 2008). Contemporary group as a fixed effect reduced bias in genetic comparisons, while when CG is considered random, the variance of the prediction error is reduced (Ramírez-Valverde *et al*., 2008).

Genomic factors

Genomic factors of bias are associated with the gene loci governing expression and are confused with environmental or residual variance (Burgueño *et al*., 2012). Genomic imprinting bias in GAS is related with production traits due to their nature as epigenetic factors (Manolio *et al*., 2009). Han *et al.* (2013) mentioned that maternal effects could be confused with genomic imprinting because they produce the same parent-of-origin patterns of phenotypic variation, leading to an over- or underestimation in GAS of traits that include maternal effects. Su *et al.* (2012) reported a bias decrease of 3.5% in genetic association values, when the additive, dominance, and epistatic effects are included in the analysis model compared to models previously reported that only included the additive effect.

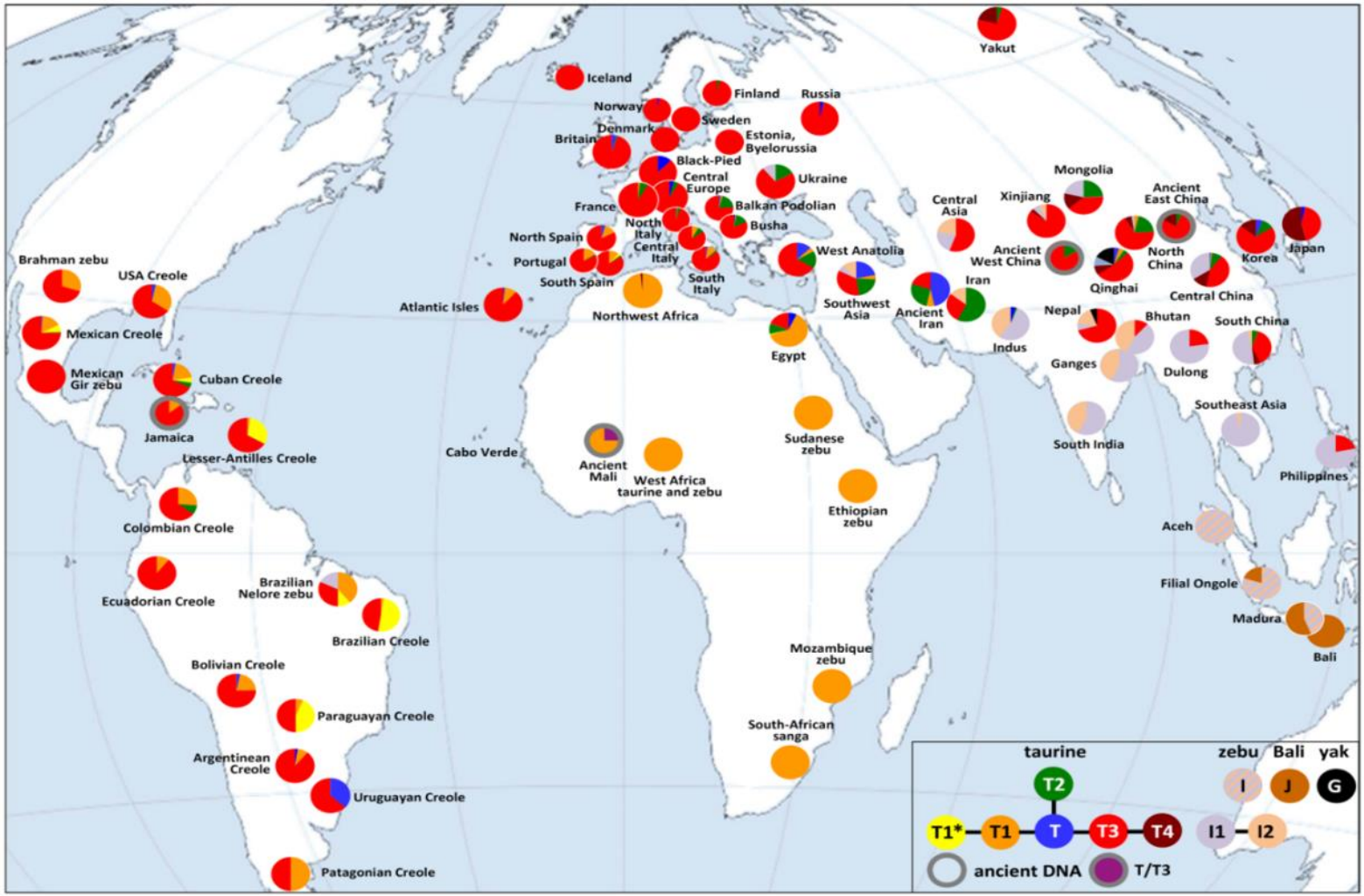

Figure 1. Diversity and distribution of major *Bos taurus* and *Bos indicus* haplogroups (taken from Lenstra *et al*., 2014).

Type of markers used in GAS is a potential source of bias due to their effect on the analysis power to determine the linkage disequilibrium (LD) level of the data (Goode and Jarvik, 2005). Additionally, Rosenberg *et al.* (2010) reported mean information content (IC) differences between microsatellites and biallelic markers across the genome, with a better performance from the second one (Figure 2). Moreover, according with Kinghorn *et al*. (2010) the correct choice of markers could increase the performance of the quantitative genotyping.

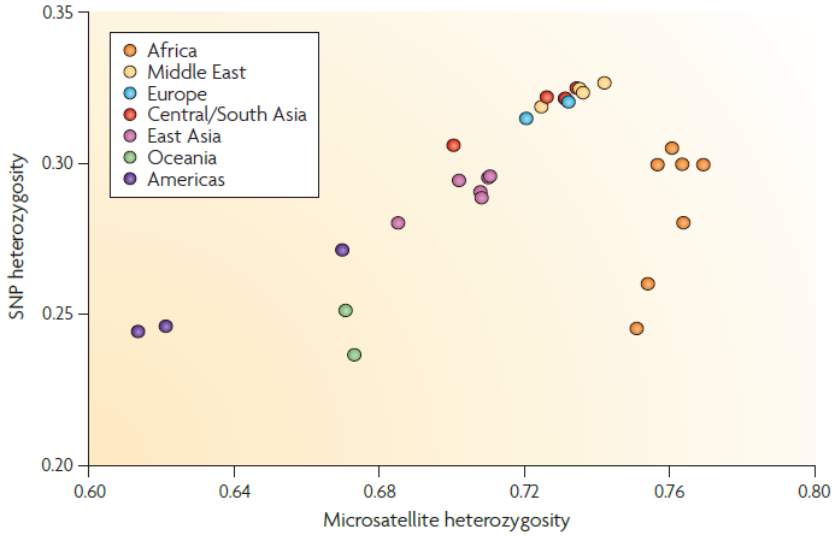

Figure 2. Information content variability for haplotype level in Europeans (taken from Rosenberg *et al*., 2010).

Monomorphism bias is based on the presence of uninformative markers in GAS (De *et al*., 2014). Thus, appearance of possible loss of power related with use of inadequate type of marker can occur. Another important genomic factor of bias is the minor allele frequency (MAF), it shows different behavior according to its effect size (Figure 3) and it is related with the Hardy-Weinberg proportions (HWP) potential bias. Therefore, MAF bias could occur if GAS use low density, monomorphic, or incorrect type of markers (Eynard *et al*., 2015).

Pleiotropic and polygenic effects are other important genetic sources of bias due to the influence over more than one economical trait in cattle (Figure 4). Pleiotropic genes, as *PLAG1*, operate like satellite regulators of the growth pathway while polygenic effect influences the estimation of genetic values.

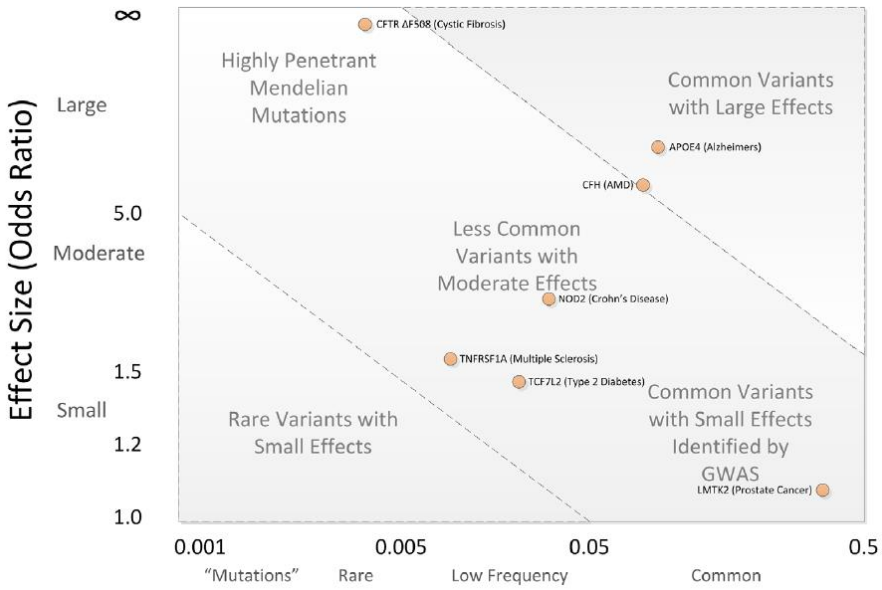

Figure 3. Types of MAF according its effect size (taken from Bush and Moore, 2012).

Segregation factor potential bias is related with the monomorphic and type of marker factors of bias and highly influences the linkage disequilibrium (LD) in the population (Bush and Moore, 2012). Since, LD describes the degree to which an allele of one SNP is inherited or correlated with the allele of another SNP within a population (De *et al*., 2014), recombination events and type of markers to detect them are critical in the development of this factor bias.

Genomic factors also include heritability bias, which is related with the gap between the phenotypic variance explained by GWAS results and those estimated by from classical heritability. Zaitlen and Kraft (2012) mentioned that "missing heritability" could be due to presence of rare variants, epistatic and gene-environment interactions, or structural variation, that are not well captured by current GWAS or their analysis methods.

Sampling-related factors

Sample selection is other source of bias. It is defined as any systematic difference between the sample and the population affecting their representativeness (Shringarpure and Xing, 2014), leading to inaccurate estimation of relationships between variables (Figure 5). According to Pyo and Wan (2012), a larger sample size is required to achieve enough

statistical power and to improve the ability of prediction. On the other hand, small sample size increases false negative rates and reduces the reliability of a study.

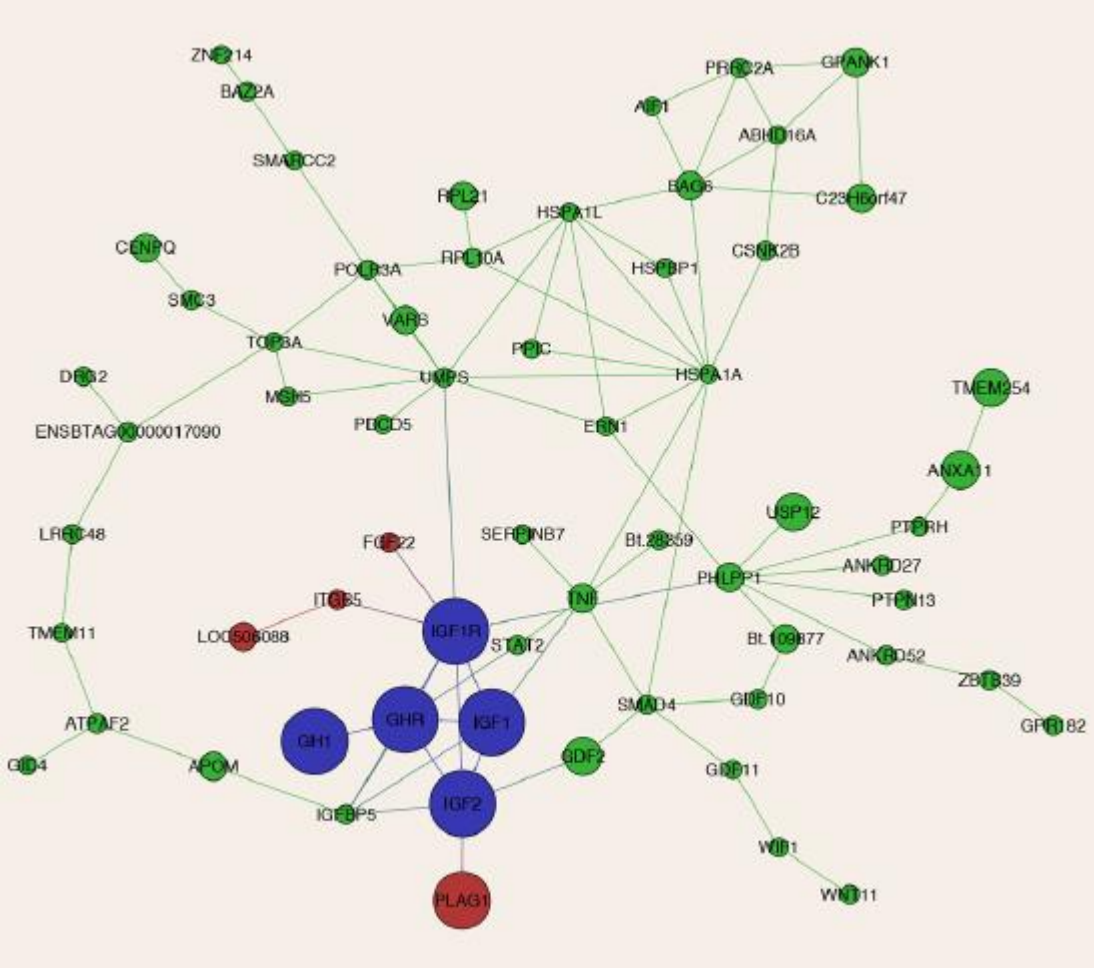

Figure 4. Network of candidate pleiotropic genes for carcass traits in Nellore cattle (taken from Pereira *et al*., 2016).

Paternity misidentification, stratification, and population structure are also factors related to sample size and its representativeness. On this regard, Visscher *et al.* (2002) determined a proportional selection response decrease of 2 to 3% for each 10% of paternity misidentification rate. Additionally, Sifuentes-Rincón *et al*. (2006) reported differences of 47% in the genetic values between simulated- and uncertain- paternity populations. Similarly, stratification bias could lead to spurious association that have no value as a tool for genetic improvement. In this sense, Zaitlen and Kraft (2012) mentioned that

stratification bias arises when there is a difference in the phenotypic variance between the population.

Statistical factors

Statistical factors of bias are those related with the model and the nature of data used. According to Pyo and Wang (2012), in genetic association studies, the observed signal for association is referred to be statistically significant if the p-value is less than a present threshold value, per example 0.05, to reject a null hypothesis if genetic association. Poor database design quality usually means high p-values and lower recognition of genetic associations (Ioannidis, 2005), especially if some of the genotypes have low frequencies in the population or traits with low heritability (Satkoski *et al*., 2011).

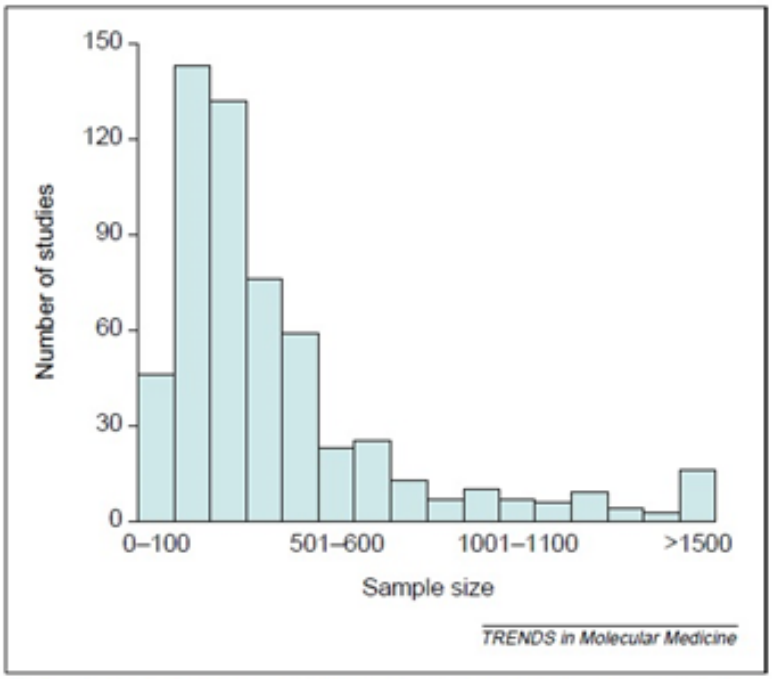

Figure 5. Sample sizes used in genetic association studies showing errors of type I (taken from Ioannidis, 2005).

Odd ratios can be a statistical factor of bias (Figure 6) when they are wrongly used as a weighted average to quantify genetic effects in GAS (Su and Lee, 2016). Due to their non-collapsible nature and tendency towards being null, a quantitative difference between

conditional and marginal odd ratios in the absence of confounding is a mathematical oddity, not a reflection of bias (Groenwold *et al*., 2011).

Another factor that could cause bias is collinearity, which refers to the non-independence of predictor variables, usually in a regression-type analysis (Dormann *et al*., 2013). Yoo *et al.* (2014) mentioned that collinearity is a problem that inflates the variance of regression parameters with a potential misidentification of relevant predictors in a statistical model. Dias *et al.* (2011) reported multicollinearity in genetic effects related with weaning weight in a Brazilian cattle population. They reported 9.8% of bias in the sum squared deviations, with variance inflation factors of 16 and 5.3 when using least square and ridge regression methodologies, respectively.

The presence of collinearity could lead to collider bias (*i.e*., the reversal paradox), an artificial association created between exposures (A and B) when a shared outcome (X) is included in the model as a covariate (Day *et al*., 2016). Day *et al.* (2016) identified over 200 spurious GAS, when the shared outcome was included as a covariate in the model used to analyze the data.

One of the most important sources of bias in GAS is the statistical model chosen due to the differences within obtained results (Figure 7). The first models used in GAS included only fixed effects, causing bias when random effects were ignored (Miciński *et al*., 2007). On the other hand, mixed models can differentiate between the effects of random error and those from systematic error (Pärna *et al*., 2012). In the same way, Maximum likelihood (ML) is another procedure used in GAS with potential of bias. Kučerová *et al.* (2006) determined that ML can estimate genetic associations of casein genes and reported mean differences in protein concentration between 42 and 73% across κ-casein genotypes (AA, AB, AE, BB, and BE). However, when estimating a higher number of associations (*e.g.*, in genome-wide association studies), the power of mixed models and ML is reduced. Extensive GAS need methods to determine the associations of thousands of markers at once. On this regard, De los Campos *et al.* (2009) reported Bayesian regression models (BM) able to adjust for the effects of thousands of markers simultaneously.

Tenesa *et al.* (2003) observed that the differences between the estimates obtained with ML and BM were small (about 5%), and both estimation procedures yielded essentially the same results. On the other hand, there are non-Bayesian models (NBM) that use

information of genotyped and non-genotyped animals to performance the genomic predictions (*e.g*. single-step genomic model) (Ma *et al*., 2015). However, due to its ability to estimate genetic association even with markers lacking information, BM and NBM are under the influence of sample size and require a pedigree as complete as possible (Sahana *et al*., 2010).

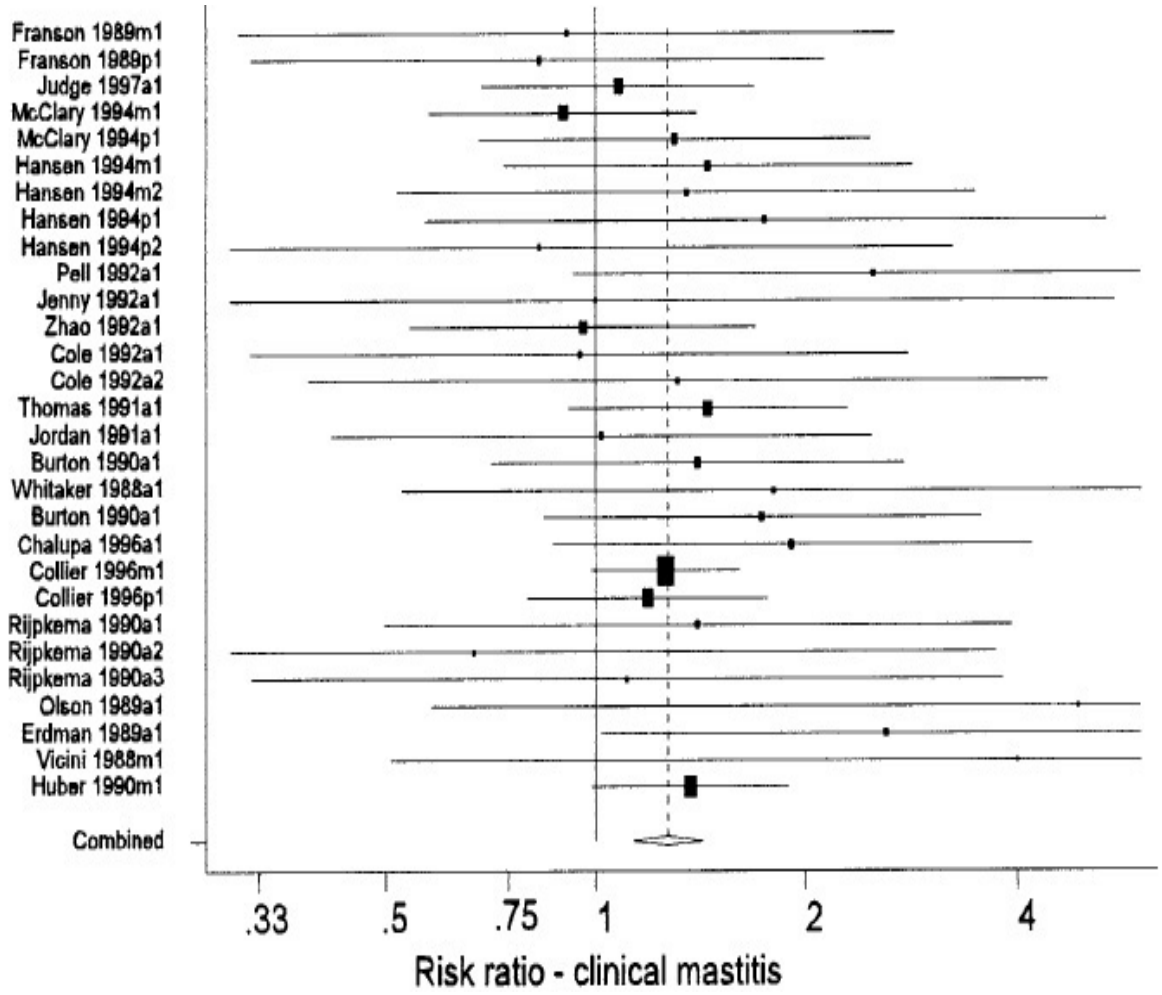

Figure 6. Forest plot of the effects or recombinant bovine somatotropin on the risk ratio of clinical mastitis (taken from Dohoo *et al*., 2003).

Strategies to correct biases in GAS

The aim of bias correction in GAS methodologies focuses on reduction of bias, rather than its elimination (Pärna *et al*., 2012). In this way, it is possible to group bias correction into genetic-genomic, statistical, and methodological strategies.

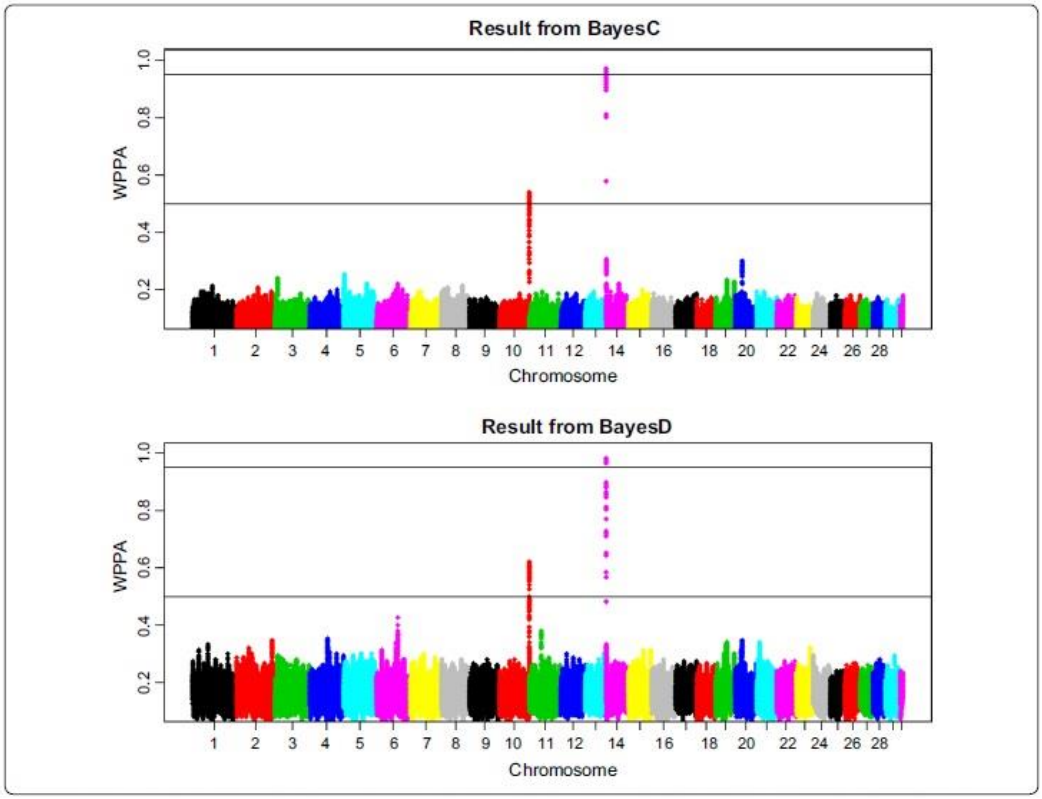

Figure 7. Probabilities of association obtained with two different Bayesian-based methods (taken from Bennewitz *et al*., 2017).

Genetic-genomic strategies

Strategies of genetic-genomic bias correction rest on two aspects: source and conditions of genetic information. The source of the genetic information in GAS refers to the approach used to obtain and report the genetic information (*i.e.*, single and multi-loci genotype or haplotype). Instead of analyzing the effects of individual alleles, some authors estimated the effects of haplotypes defined by genes associated with the traits under study (Zhou *et al*., 2013), while other authors used multi-loci genotypes for the same purpose (Jaiswal *et al*., 2016).

The use of haplotypes and multi-loci genotypes can reduce bias arising from the way several genes are combined, the polygenic effect of the studied traits, and the position of the analyzed loci within the genome. However, unlike multi-loci genotypes, it has been argued that haplotypes have similar effects on different breeds (Andrés *et al*., 2007). As a

result, a common approach to analyzing the effects of haplotype has been to determine the most likely configuration for each and assume that this allocation of haplotypes is known without error when subsequent statistical analyses are performed. However, precise haplotype construction could be difficult, and often leads to biased estimates and reduced analytical power in GAS (Andrés *et al.,* 2007). In addition, when multiple loci are genotyped, haplotypes are unknown because there is no information about linkage phase of alleles at different loci (Sahana *et al*., 2010). Sahana *et al.* (2010) observed a high rate of type I error when using haplotypes as a fixed effect in genetic association models. Zhang *et al.* (2016) concluded that when there is a lack of tools available to reconstruct haplotypes, the best alternative is to use multi-loci genotypes regardless of whether phase adjustment information is available.

Other factors affecting the reliability of results are the number of markers used for reconstruction and the way that haplotypes and multi-loci genotypes are included in GAS models. For reconstruction, the best results have been obtained using 2 to 5 markers (Abdallah *et al*., 2004). In this sense, the main benefit of using haplotypes or multi-loci genotypes is their ability to explain most of the additive, dominance, and epistasis effects on the loci studied (Zhao *et al*., 2012). With respect to inclusion methods, incorporating haplotype as a random effect imparts better performance compared with models that include it as a fixed effect in terms of power, control of type I error, and precision (Boleckova *et al*., 2012). Hence, some of the probable HWP bias in these studies can be avoided, especially if the nature of the alleles being studied is considered. Kent *et al.* (2007) concluded that due to the risk of wrong associations, it is best to use common genetic variants greater than 10% as rare alleles generate biases in their association values and equally affect the values of common alleles. Therefore, the conditions needed to establish the use of haplotypes, genotypes, or both in GAS are of utmost importance for devising strategies to correct bias of genetic information.

Sampling-related and statistical strategies

Methodological strategies used to avoid sampling bias are based on grouping of individuals or samples that share the same features in order to reduce heterogeneity and increase the representativeness of the results (Gustavsson *et al*., 2014). On the other hand, the use of previously reported information more important when establishing a methodological bias reduction strategy. Published information enables to use features and results previously validated and helps to avoid the risk of bias related with transferring results among breeds (Poulsen *et al*., 2015).

Methodological strategies to reduce bias associated with the statistical source are based on reviews, as well as the use of estimates and other results in the literature to determine the best models and features for the phenomenon studied (Brito *et al*., 2011). Association methods commonly used are based on family structure (pedigree) and case-control studies with unrelated individuals (De los Campos *et al*., 2009). However, case-control studies are the most viable to study genetic association because studies based on family structure involves extended periods of testing (Kent *et al*., 2007). The presence of type I errors due to the subjective nature of the estimates (underlying assumptions) could address the risk of under- or overestimation of studied traits (Zoche-Golob *et al*., 2015). Therefore, the best strategy to reduce statistical bias lies in all aspects related to the predictive power of the approaches since it depends on all elements of bias that might arise.

In conclusion, it is necessary to consider the methodology used in previous GAS to establish a hierarchy of sources of bias and to facilitate better decisions on the use of tools to reduce inconsistencies in the results of future studies.

Acknowledgments

Authors thank to CONACY for the financial support to the first and third authors during their doctoral studies.

Conflicts of interest

The authors declare they have no conflicts of interest with regard to the work presented in this report.

References

Abdallah JM, Mangin B, Goffinet B, Cierco-Ayrolles C, Perez-Enciso M. A comparison between methods for linkage disequilibrium fine mapping of quantitative trait loci. Genet Res 2004; 83: 41-47. https://www.researchgate.net/publication/8580615

Andrés AM, Clark AG, Shimmin L, Boerwinkle E, Sing CF, Hixson JE. Understanding the accuracy of statistical haplotype inference with sequence data of known phase. Genet Epidemiol 2007; 31: 659-671. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2291540/pdf/nihms-44071.pdf

Bennewitz J, Edel C, Fries R, Meuwissen THE, Wellmann R. Application of a Bayesian dominance model improves power in quantitative trait genome-wide association analysis. Genet Sel Evol 2017; 49: 7. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5237573/pdf/12711_2017_Article_284.pdf

Benyamin B, Visscher PM, McRae A. Family-based genome-wide association studies. Pharmacogenomics 2009; 10(2): 181-190. https://www.futuremedicine.com/doi/10.2217/14622416.10.2.181

Boleckova J, Christensen OF, Sørensen P, Sahana G. Strategies for haplotype-based association mapping in a complex pedigreed population. Czech J Anim Sci 2012; 57: 1-9. http://pure.au.dk/portal/files/44004867/Jana_Haplotype.pdf

Brito LF, Silva FG, Melo ALP, Caetano GC, Torres RA, Rodrigues MT, Menezes GRO. Genetic and environmental factors that influence production and quality of milk of Alpine and Saanen goats. Genet Mol Res 2011; 10: 3794-3802. http://www.geneticsmr.com/sites/default/files/articles/year2011/vol10-4/pdf/gmr1505.pdf

Burgueño J, De los Campos G, Weigel K, Crossa J. Genomic prediction of breeding values when modeling genotype × environment interaction using pedigree and dense molecular markers. Crop Science 2012; 52: 707–719. http://dysci.wisc.edu/wp-content/uploads/sites/40/2013/11/cs-52-2-707.pdf

Bush WS, Moore JH. Chapter 11: Genome-Wide Association Studies. PLoS Comput Biol 2012; 8(12): e1002822. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3531285/pdf/pcbi.1002822.pdf

Day FR, Loh PR, Scott RA, Ong KK, Perry JRB. A robust example of collider bias in a genetic association study. Am J Hum Genet 2016; 98: 392–393. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4746366/pdf/main.pdf

De R, Bush WS, Moore JH. Bioinformatics challenges in genome-wide association studies (GWAS). In: Ronald Trent editor. Clinical Bioinformatics, Methods in Molecular Biology. New York: Springer Science Bussiness Media; 2014. p.63-81.

De los Campos, G., Naya H., Gianola D., Crossa J., Legarra A., Manfredi E., Weigel K., Cotes J. M. 2009. Predicting quantitative traits with regression models for dense molecular markers and pedigree. Genetics 182:375-385. http://www.genetics.org/content/genetics/182/1/375.full.pdf

Deb R, Singh U, Kumar S, Singh R, Sengar G, Sharma A. Genetic polymorphism and association of kappa-casein gene with milk production traits among Frieswal (HF x Sahiwal) cross breed of Indian origin. Iran J Vet Res 2014; 15(4): 406-408. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4789222/pdf/ijvr-15-406.pdf

Dias RAP, Petrini J, Sterman JB, Pereira J, Santos R, Lopes AL, Barreto G. Multicollinearity in genetic effects for weaning weight in a beef cattle composite population. Livest Sci 2011; 142: 188–194. http://www.sciencedirect.com/science/article/pii/S1871141311002836

Dickerson GE. Evaluation of breeds and crosses of domestic animals. Rome(IT): Animal Production and Health Paper FAO; 1993.

Dohoo IR, DesCoteaux L, Leslie K, Fredeen A, Shewfelt W, Preston A, Dowling P. A meta-analysis review of the effects of recombinant bovine somatotropin 2. Effects on animal health, reproductive performance, and culling. Can J Vet Res 2003; 67: 252-264. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC280709/pdf/20031000s00002p252.pdf

Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, Carré G, García JR, Gruber B, Lafourcade B, Leitão PJ, Münkemüller T, McClean C, Osborne PE, Reineking B, Schröder B, Skidmore AK, Zurell D, Lautenbach S. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 2013; 36: 027–046. http://onlinelibrary.wiley.com/doi/10.1111/j.1600-0587.2012.07348.x/epdf

Duifhuis-Rivera T, Lemus-Flores C, Ayala-Valdovinos MÁ, Sánchez-Chiprés DR, Galindo-García J, Mejía-Martínez K, González-Covarrubias E. Polymorphisms in beta and kappa casein are not associated with milk production in two highly technified populations of Holstein cattle in México. J Anim Plant Sci 2014; 24: 1316-1321. http://www.thejaps.org.pk/docs/v-24-5/07.pdf

Eynard SE, Windig JJ, Leroy G, Van Binsbergen R, Calus MPL. The effect of rare alleles on estimated genomic relationships from whole genome sequence data. BMC Genetics 2015; 16: 24. http://bmcgenet.biomedcentral.com/articles/10.1186/s12863-015-0185-0

Foulkes AS. Applied statistical genetics with R for population-based association studies. Cham(Swiss): Springer International Publishing AG; 2009.

Goode EL, Jarvik GP. Assessment and implications of linkage disequilibrium in genome-wide single-nucleotide polymorphism and microsatellite panels. Genet Epidemiol 2005; 29(Suppl.1): S72-S76. http://onlinelibrary.wiley.com/doi/10.1002/gepi.20112/pdf

Groenwold RHH, Moons KGM, Peelen LM, Knol MJ, Hoes AW. Reporting of treatment effects from randomized trials: A plea for multivariable risk ratios. Contemp Clin Trials 2011; 32: 399-402. http://www.contemporaryclinicaltrials.com/article/S1551-7144(10)00247-8/pdf

Gustavsson F, Buitenhuis AJ, Johansson M, Bertelsen HP, Glantz M, Poulsen NA, Lindmark-Månsson H, Stålhammar H, Larsen LB, Bendixen C, Paulsson M, Andrén A. Effects of breed and casein genetic variants on protein profile in milk from Swedish Red, Danish Holstein, and Danish Jersey cows. J Dairy Sci 2014; 97: 3866-3877. http://www.journalofdairyscience.org/article/S0022-0302(14)00236-7/pdf

Han M, Hu YQ, Lin S. Joint detection of association, imprinting and maternal effects using all children and their parents. Eur J Hum Genet 2013; 21: 1449–1456. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3831068/pdf/ejhg201349a.pdf

Ioannidis JP. Why most published research findings are false. Plos Med 2005; 2: e124. http://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.0020124

Jaiswal V, Gahlaut V, Meher PK, Mir RR, Jaiswal JP, Rao AR, Bayan HS, Gupta PK. Genome wide single locus single trait, multi-locus and multi-trait association mapping for some important agronomic traits in common wheat (*T. aestivum L.*). Plos One 2016; 11(7): e0159343. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4956103/pdf/pone.0159343.pdf

Jahuey-Martínez FJ, Parra-Bracamonte GM, Sifuentes-Rincón AM, Martínez-González JC, Gondro C, García-Pérez CA, López-Bustamantes LA. Genomewide association

analysis of growth traits in Charolais beef cattle. J Anim Sci 2016; 94: 4570-4582. https://www.ncbi.nlm.nih.gov/pubmed/27898967

Kent JW, Dyer TD, Göring HH, Blangero J. Type I error rates in association versus joint linkage/association tests in related individuals. Genet Epidemiol 2007; 31: 173-177. http://onlinelibrary.wiley.com/doi/10.1002/gepi.20200/epdf

Kinghorn BP, Bastiaansen JWM, Ciobanu DC, Van Der Steer HAM. Quantitative genotyping to estimate genetic contributions to pooled samples and genetic merit of the contributing entities. Acta Agric Scand A 2010; 60: 3-12. http://www.tandfonline.com/doi/abs/10.1080/09064701003801922

Kučerová J, Matějiček A, Jandurová OM, Sørensen P, Němcová E, Štípková M, Kott V, Bouška J, Frelich J. Milk protein genes CSN1S1, CSN2, CSN3, LGB and their relation to genetic values of milk production parameters in Czech Fleckvieh. Czech J Anim Sci 2006; 51: 241-247. http://agriculturejournals.cz/publicFiles/52288.pdf

Lee YH. Meta-analysis of genetic association studies. Ann Lab Med 2015; 35: 283-287. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4390695/pdf/alm-35-283.pdf

Lenstra JA, Ajmone-Marsan P, Beja-Pereira A, Bollongino R, Bradley DG, Colli L, De Gaetano A, Edwards CJ, Felius M, Ferretti L, Ginja C, Hristov P, Kantanen J, Lewis CM, Knight J. Introduction to genetic association studies. Cold Spring Harbor Protoc 2012; 3: 297-306. http://cshprotocols.cshlp.org/cgi/pmidlookup?view=long&pmid=22383645

Lirón JP, Magee DA, Negrini R, Radoslavov GA. Meta-analysis of mitochondrial DNA reveals several population bottlenecks during worldwide migrations of cattle. Diversity 2014; 6: 179-187. http://www.mdpi.com/1424-2818/6/1/178

Ma P, Lund MS, Nielsen US, Aamand GP, Su G. Single-step genomic model improved reliability and reduced the bias of genomic predictions in Danish Jersey. J Dairy Sci 2015; 98: 9026-9034. http://www.sciencedirect.com/science/article/pii/S0022030215006992?via%3Dihub

Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos ME, Cardon LR, Chakravarti A, Cho JH, Guttmacher AE, Kong A, Kruglyak L, Mardis E, Rotimi CE, Slatkin M, Valle D, Whittemore AS, Boehnke M, Clark AG, Eichler EE, Gibson G, Haines JL, Mackay TFC, McCarroll SA, Visscher PM. Finding the missing heritability of complex diseases. Nature 2009; 461(7265): 747–753. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2831613/

Miciński J, Klupczyński J, Mordas W, Zablotna R. Yield and composition of milk from Jersey cows as dependent on the genetic variants of milk proteins. Pol J Food Nutr Sci 2007; 57: 95-99. http://journal.pan.olsztyn.pl/pdfy/2007/3A/21.pdf

Pärna E, Kaart T, Kiiman H, Bulitko T, Viinalass H. In: Chaiyabutr N, editor. Milk Production-Advanced Genetic Traits, Cellular Mechanism, Animal Management and Health. Rijeka: InTech; 2012. p.155-172.

Pereira AGT, Utsunomiya YT, Milanesi M, Torrecilha RBP, Carmo AS, Neves HHR, Carvalheiro R, Ajmone-Marsan P, Sonstegard TS, Sölkner J, Contreras-Castillo CJ, Garcia JF. Pleiotropic genes affecting carcass traits in *Bos indicus* (Nellore) cattle are modulators of growth. Plos One 2016; 11(7): e0158165. http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0158165

Poulsen NA, Buitenhuis AJ, Larsen LB. Phenotypic and genetic associations of milk traits with milk coagulation properties. J Dairy Sci 2015; 98: 1-9. http://www.journalofdairyscience.org/article/S0022-0302(15)00084-3/pdf

Pyo HE, Wan PJ. Sample size and statistical power calculation in genetic association studies. Genomics Inform 2012; 10(2): 117-122. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3480678/

Ramírez-Valverde R, Núñez-Domínguez R, Ruíz-Flores A, García-Muñiz JG, Magaña-Valencia F. Comparison of contemporary group definitions for genetic evaluation of Braunvieh cattle. Téc Pec Mex 2008; 46(4): 359-370. http://cienciaspecuarias.inifap.gob.mx/index.php/Pecuarias/article/viewFile/1799/1793

Rosenberg NA, Huang L, Jewett EM, Szpiech ZA, Jankovic I, Boehnke M. Genome-wide association studies in diverse populations. Nat Rev Genet 2010; 11(5): 356-366. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3079573/pdf/nihms279798.pdf

Sagoo GS, Little J, Higgins JPT. Systematic reviews of genetic association studies. PLoS Med 2009; 6(3): e1000028. http://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1000028

Sahana G, Guldbrandtsen B, Janss L, Lund MS. Comparison of association mapping methods in a complex pedigreed population. Genet Epidemiol 2010; 34: 455–462. http://onlinelibrary.wiley.com/doi/10.1002/gepi.20499/pdf

Satkoski JA, Malhi RS, Kanthaswamy S, Johnson J, Garnica WT, Malladi VS, Smith DG. The effect of SNP discovery method and sample size on estimation of population genetic data for Chinese and Indian rhesus macaques (*Macaca mulatta*). Primates 2011; 52: 129-138. https://www.ncbi.nlm.nih.gov/pubmed/21207104

Schwarzenbacher H, Burgstaller J, Seefried FR, Wurmser C, Hilbe M, Jung S, Fuerst C, Dinhopl N, Weissenböck H, Fuerst-Waltl B, Dolezal M, Winkler R, Grueter O, Bleu U, Wittek T, Fries R, Pausch H. A missense mutation in TUBD1 is associated with high juvenile mortality in Braunvieh and Fleckvieh cattle. BMC Genomics 2016; 17: 400. https://bmcgenomics.biomedcentral.com/articles/10.1186/s12864-016-2742-y

Shringarpure S, Xing EP. Effects of sample selection bias on the accuracy of population structure and ancestry inference. G3-Genes Genom Genet 2014; 4: 901-911. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4025489/

Sifuentes-Rincón AM, Parra-Bracamonte GM, De la Rosa RXF, Sánchez VA, Serrano MF, Rosales AJ. Importancia de las pruebas de paternidad basadas en microsatélites para la evaluación genética de ganado de carne en empadre múltiple. Tec Pec Mex 2006; 44(3): 389-398. http://cienciaspecuarias.inifap.gob.mx/index.php/Pecuarias/article/view/1732

Su YS, Lee WC. False appearance of gene–environment interactions in genetic association studies. Medicine 2016; 95(9): 1-8. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4782844/

Su G, Christensen OF, Ostersen T, Henryon M, Lund MS. Estimating additive and non-additive genetic variances and predicting genetic merits using genome-wide dense single nucleotide polymorphism markers. Plos One, 2012, 7(9): e45293. http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0045293

Tenesa A, Knott SA, Ward D, Smith D, Williams JL, Visscher PM. Estimation of linkage disequilibrium in a sample of the United Kingdom dairy cattle population using unphased genotypes. J Anim Sci 2003; 81: 617-623. https://keppel.qimr.edu.au/contents/p/staff/CVPV066.pdf

Trail JCM, Gregory KE, Durkin J, Sandford J. Crossbreeding cattle in beef production programmes in Kenya. II. Comparison of purebred Boran and Boran crossed with the Red Poll and Santa Gertrudis breeds. Trop Anim Hlth Prod 1984; 16: 191-200. https://link.springer.com/article/10.1007/BF02265318

Visscher PM, Woolliams JA, Smith D, Williams JL. Estimation of pedigree errors in the UK dairy population using microsatellite markers and the impact on selection. J Dairy Sci 2002; 85: 2368–2375.
http://www.sciencedirect.com/science/article/pii/S0022030202743178

Wu C, Li S, Cui Y. Genetic Association Studies: An Information Content Perspective. Curr Genomics 2012; 13(7): 566-573.
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3468889/

Yoo W, Mayberry R, Bae S, Singh K, He Q, Lillard J. A study of effects of multicollinearity in the multivariable analysis. Int J Appl Sci Technol 2014; 4(5): 9–19.
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4318006/pdf/nihms653352.pdf

Zaitlen N, Kraft P. Heritability in the genome-wide association era. Hum Genet 2012; 131(10): 1655-1664. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3432754/

Zhang W, Li J, Guo Y, Zhang L, Xu L, Gao X, Zhu B, Gao H, Ni H, Chen Y. Multi-strategy genome-wide association studies identify the DCAF16-NCAPG region as a susceptibility locus for average daily gain in cattle. Sci Rep 2016; 6: 38073.
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5125095/

Zhao H, Rebbeck TR, Mitra N. Analyzing genetic association studies with an extended propensity score approach. Stat Appl Genet Mol Biol 2012; 11(5): 1-17.
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3518898/

Zhou L, Ding X, Zhang Q, Wang Y, Lund MS, Su G. Consistency of linkage disequilibrium between Chinese and Nordic Holsteins and genomic prediction for Chinese Holsteins using a joint reference population. Genet Sel Evol 2013; 45: 1-7.
http://.www.gsejournal.biomedcentral.com/articles/10.1186/1297-9686-45-7

Zoche-Golob V, Heuwieser W, Krömker V. Investigation of the association between the test day milk fat-protein ratio and clinical mastitis using a Poisson regression approach for analysis of time-to-event data. Prev Vet Med 2015; 121: 64-73. http://www.sciencedirect.com/science/article/pii/S016758771500230

3. POTENTIAL INFLUENCE OF CSN3 AND LGB GENES IN GENETIC ASSOCIATION STUDIES OF MILK QUALITY TRAITS

José Luis Zepeda-Batista, Luis Antonio Saavedra-Jiménez, Agustín Ruíz-Flores*, Rafael Núñez-Domínguez, Rodolfo Ramírez-Valverde

Departamento de Zootecnia, Posgrado en Producción Animal, Universidad Autónoma Chapingo, Km 38.5 Carretera México-Texcoco Chapingo, México 56230.

Published in Asian-Australasian Journal of Animal Science 30(2):1684-1688.

Doi: http://dx.doi.org/10.5713/ajas.16.0481

Running title: Influence of CSN3 and LGB genes on milk traits

Potential influence of CSN3 and LGB genes in genetic association studies of milk quality traits

José Luis Zepeda-Batista, Luis Antonio Saavedra-Jiménez, Agustín Ruíz-Flores*, Rafael Núñez-Domínguez, Rodolfo Ramírez-Valverde

Posgrado en Producción Animal, Universidad Autónoma Chapingo, Chapingo, Estado de México, Zip Code 56230, México. E-mail: produccionanimal@correo.chapingo.mx

*Author for correspondence: Agustín Ruíz-Flores, E-mail: arf@correo.chapingo.mx, Phone/Fax: +52 595 952 1621.

ABSTRACT

Objective: From a review of published information on genetic association studies, a meta-analysis was conducted to determine the influence of the genes CSN3 (κ-casein) and LGB (β-lactoglobulin) on milk yield traits in Holstein, Jersey, Brown Swiss, and Fleckvieh.

Methods: The GLIMMIX procedure was used to analyze milk production and percentage of milk protein and fat by trait. Models included the main effects and all their possible two-way interactions; not estimable effects and non-significant ($p>0.05$) two-way interactions were dropped from the models. The three traits analyzed used Poisson distribution and a log link function and were determined with the Interactive Data Analysis of SAS software. Least square means and multiple mean comparisons were obtained and performed for significant main effects and their interactions ($p<0.0255$).

Results: Interaction of breed by gene showed that Holstein and Fleckvieh were the breeds on which CSN3 (6.01±0.19 and 5.98±0.22%), and LGB (6.02±0.19 and 5.70±0.22%) have the greatest influence. Interaction of breed by genotype nested in the analyzed gene indicated that Holstein and Jersey showed greater influence of the CSN3 AA genotype, 6.04±0.22 and 5.59±0.31% than the other genotypes, while LGB AA genotype had the largest influence on the traits analyzed, 6.05±0.20 and 5.60±0.19%, respectively. Furthermore, interaction of type of statistical model by genotype nested in the analyzed gene indicated that CSN3 and LGB genes had similar behavior, maintaining a difference of more than 7% across analyzed genotypes. These results could indicate that both Holstein and Jersey have had lower substitution allele effect in selection programs that include CSN3 and LGB genes than Brown Swiss and Fleckvieh.

Conclusion: Breed determined which genotypes had the greatest association with analyzed traits. The mixed model based in Bayesian or Ridge Regression was the best alternative to analyze CSN3 and LGB gene effects on milk yield and protein and fat percentages.

Keywords: Dairy Cattle; Genetic Improvement; Polymorphism

INTRODUCTION

Genetic association studies have been increasingly used in cattle breeding programs. However, the results have been inconsistent for milk protein genes. Positive, negative, or absence of association with similar genotypes have been reported [1]. Examples of these are the conclusions of Duifhuis-Rivera et al. [2] and Dogru [3] who reported that different CSN3 (□-casein) and LGB (□-lactoglobulin) genotypes were not associated with milk yield in a Mexican herd of Holstein and in a Turkish herd of Brown Swiss cattle, respectively. On the contrary, Gustavsson et al. [4] reported that the composite genotype BB/A1A2/AB of CSN1S1/CSN2/CSN3 has positive effects on cheese yield and percentage of protein and fat in milk.

The inconsistencies have been attributed to various issues affecting the production and composition of milk. Bernabucci et al. [5] concluded that the sampling station affects total protein concentration and acidity in milk, ranging from 3.2 to 6.0% and from 2.0 to 5.6%, respectively. Moreover, Streit et al. [6] documented the importance of determining the alleles in the DGAT1 genotype in German Holstein sires to avoid differences that can be as high as 20% in the association of these alleles with milk production values as a result of the allele substitution effect.

However, there is no accurate information related to the potential influence of CSN3 and LGB genes that could be used in genetic association studies to analyze samples obtained under different racial conditions and sampling methodologies. Therefore, in this study we conducted a meta-analysis of genetic association studies to determine the behavior of genes CSN3 (Genbank Accession No. AC_000163.1) and LGB (Genbank Accession No. AC_000168.1) on milk yield and percentage of protein and fat under different conditions and methods of analysis.

MATERIALS AND METHODS

Paper selection criteria

Scientific journals were searched for published papers on genetic association studies. Initially, one hundred and forty-seven papers were chosen. The used criteria aimed to eliminate their heterogeneity, to look for representability of the results, and to ensure their replicability. In a first step, the papers published from 2003 to 2016 were selected. Subsequently, we considered only studies dealing with the most studied milk protein

genes, CSN3 and LGB. The remaining papers described studies on milk production and protein and fat percentage in milk. The final sample included 26 papers dealing with Holstein, Jersey, Brown Swiss, and Fleckvieh dairy breeds.

Meta-analysis

Traits analyzed were milk production and percentage of milk protein and fat. The random and fixed effects included in the final model were determined by first establishing a complete model by trait. This model included the main effects and all their possible two-way interactions; some of these were not estimable and were dropped from the model. Additionally, non-significant ($p>0.05$) two-way interactions were deleted as well. The best fit model and link function were determined with the modules Distribution (Y) and Fit (Y, X), respectively; both are modules from the Interactive Data Analysis of SAS software [7]. The three traits analyzed had the best fit with a final model that used Poisson distribution and a log link function. The final models for milk production, protein and fat percentages were as follows:

$$MP = A_i + B_z + M_y + T_x + AG_w + GN_v + SS_u + BAG_t + BGN_s + MGN_r + E$$

where MP= milk production; Ai= random effect of i-th paper included in the study; Bz= z-th effect of the breed; My= y-th effect of the used model in the original study; Tx= x-th effect of the used test to recover the original data; AGw= w-th effect of the analyzed gene; GNV= v-th effect of the genotype nested in the analyzed gene; SSu= u-th effect of the sample size used; BAGt= t-th effect of the interaction of breed by analyzed gene; BGNs= s-th effect of the interaction of breed by genotype nested in the analyzed gene; MGNr= r-th effect of the interaction model by genotype nested in the analyzed gene; and E is the residual random effect.

$$P = A_i + B_z + M_y + T_x + AG_w + GN_v + SS_u + E$$

where P= protein percentage; Ai= random effect of i-th paper included in the study; Bz= z-th effect of the breed; My= y-th effect of the used model in the original study; Tx= x-th effect of the used test to recover the original data; AGw= w-th effect of the analyzed gene; GNV= v-th effect of the genotype nested in the analyzed gene; SSu= u-th effect of the sample size used; and E is the residual random effect.

$$F = A_i + B_z + M_y + T_x + AG_w + GN_v + SS_u + E$$

where F= fat percentage; Ai= random effect of i-th paper included in the study; Bz= z-th effect of the breed; My= y-th effect of the used model in the original study; Tx= x-th effect of the used test to recover the original data; AGw= w-th effect of the analyzed gene; GNV= v-th effect of the genotype nested in the analyzed gene; SSu= u-th effect of the sample size used; and E is the residual random effect.

Once all the information from the papers of the sample had been gathered, an adjustment of the information expressed as deviations from their mean was conducted. The population mean reported in each article was used as a reference mean for the analyzed traits. The adjusted values were included in a final database and the statistical analysis was conducted. The GLIMMIX procedure [7] was used to analyze the information. For the main effects and their interactions that were significant ($p<0.0255$), least square means were obtained and multiple mean comparisons were performed.

The main effects considered in the analysis were the following: article (A), each article included in the study; breed (B), Holstein, Jersey, Brown Swiss, and Fleckvieh; model (M), each of the statistical analysis approaches used in the papers (least square means and mixed models based on Bayesian or Ridge Regression: MMBRR); test (T), the test to recover the data used in the papers (305-day and one-day test); gene (AG), CSN3 and LGB; genotype (GN), AA, AB, AE, BB, BE, and BE for CSN3, and for LGB were AA, AB, and BB; sample size (SS), size of the population analyzed in the paper, a) 1 to 500 animals was considered small, and b) from 501 to 2000 was regarded as medium. The final model for milk production included some two-way interactions: breed by gene (BAG); breed by genotype nested in analyzed gene (BGN); and model by genotype nested in analyzed gene (MGN). The effect of paper was regarded as random, the rest of them were considered fixed.

RESULTS AND DISCUSSION

Table 1 presents the level of significance for the effects considered in the models to analyze milk production (MP) and protein (PP) and fat percent (FP) in milk. The type of test used was highly significant ($P < 0.0001$) for milk production, probably because the data used in the meta-analysis included one-day tests as well as 305-day tests. On this regard, Gustavsson et al. [4] and Poulsen et al. [8] concluded that the biases associated with these tests could be controlled by grouping measurements obtained with the same test.

The sample size was significant ($p<0.0001$) for milk production. On this regard, Vidović et al. [9] concluded that bias caused by the sample size is due to the difficulty of differentiating the effects of each gene in the analysis of polygenic traits. On the other hand, the effect of the interaction model by genotype nested in the analyzed gene was significant ($p\leq0.0255$) probably because of the variance explained by the random effects considering the mixed model [9,10]

The interaction of breed by analyzed gene, was highly significant ($p<0.0001$). The effect of the gene analyzed had different behavior across breeds and between genes within breed (Table 2). The largest difference in CSN3 and LGB was between the Holstein and Jersey breeds, with 7.65% and 7.64% higher in Holstein than in Jersey, respectively. In Fleckvieh and Brown Swiss breeds, both genes are similarly associated with milk production and total solids, showing greater influence of both genes in Fleckvieh, up to 7.65% and 7.64% for CSN3 and LGB, respectively, compared with Brown Swiss.

Deb et al. [11], mentioned that breed has a marked effect on milk yield but somewhat less of an effect on milk composition. The heavier breeds tend to produce more milk. Therefore, the breeds with high milk production (Holstein and Fleckvieh) showed the greatest response. On this regard, since the improvement programs in cattle are based on the genetic potential that the individuals could show. These results may be due to selection for traits such as milk production and total solids that are associated with CSN3 and LGB in both breeds [12,13]. Additionally, genetic make-up of dairy animals plays a great role in the variation of milk yield and composition. Raven et al. [14] and Ramayo-Caldas et al. [15] mentioned the importance of determine the proportion of genetic markers shared between breeds to study and compare them in multi-breed multi-trait association studies in commercial herds. On this regard, Ramayo-Caldas et al. [15], reported around 206 genes with the same effect in three French breeds. Meanwhile, Raven et al. [14], determined that despite the different linkage disequilibrium patterns, Holstein and Jersey share between 8% to more than 38% genetic markers with similar effects on economic important traits in dairy cattle.

The interaction of breed by genotype nested in the analyzed gene (Table 3) was highly significant ($p<0.0001$). Holstein and Jersey had a greater influence of genotype AA in CSN3 gene on milk yield, up 1.61% from other genotypes studied; Miciński et al. [12] reported a difference of 0.92% to 11.94% for Polish Jersey.

Estimates for Brown Swiss and Fleckvieh were different from those for Holstein and Jersey; the BB genotype was the most closely associated with milk production, 7.94% higher than the other studied genotypes. Similarly, Chrenek et al. [16] found that, for Fleckvieh, the genotype most closely associated with milk yield is BB, 7.42%. In contrast, Matějiček et al. [17] concluded that genotype AB was the most closely associated with that trait.

Similar behavior shown in the study by Holstein and Jersey and Brown Swiss and Fleckvieh may also attend mainly the common geographic origin of the analyzed breeds. In this regard, Negrini et al. [18] calculated genetic distances through genetic fingerprinting of 51 cattle breeds. They placed the Holstein and Jersey breeds in the Nordic genetic type group, while the Brown Swiss and Fleckvieh were grouped with genetic types from the Alpine region or Central-France.

Holstein and Jersey showed similar behavior for AA genotype of the LGB gene (Table 3). This estimate had the largest influence on milk yield, up to 0.66% and 1.07% above the other genotypes. This result is similar to that found for the CSN3 gene in this study since, according to Bonfatti et al. [19], the LGB gene is also associated with milk production and total solids. Fleckvieh and Brown Swiss are influenced in a major way by genotypes BB and AB, showing up to 1.77% and 1.05% difference, relative to other genotypes. On this regard, Gustavsson et al. [4] determined the effect of breed and genotype using composite genotypes of □□□□casein and β-lactoglobulin genotypes to determine their genetic association with production traits. These authors concluded that the differences between the genetic association values, ranging between 0.5% and 30%, were mainly influenced by genotype. Moreover, Oltenacu and Broom [20], report rising of inbreeding rates of 0.2% per year in Holstein and Jersey which would cause a decreasing in the response to improvement and selection program and the genetic association values due to loss of the genetic variation in the population.

On the other hand, the interaction of type of statistical model used by genotype nested in the analyzed gene was highly significant ($p<0.0001$), and mean estimates are shown in Table 4.

The differences between the models used for each of the studied genes may be explained with the conclusions of Miciński et al. [12] and Monir and Zhu [21], who demonstrated that the inclusion or omission of the effect of each particular gene and polygenic effects, influence the results of production traits and only the models that include random effects are able to differentiate those changes.

The results of comparing models of LGB and CSN3 genes (Table 4) were similar, maintaining a difference of just over 7% across all genotypes. Here, Comin et al. [10] and Vidović et al. [9] concluded that the MMBRR explained the effect of milk protein genotypes on performance and composition of different breeds of dairy cattle, nearing a 9% difference, relative to studies using only models with fixed effects based on least squares.

Oleński et al. [22] and Pärna et al. [23] used models similar to ours when attempted to explain the variability present in genetic association studies. However, the general linear mixed models were not able to differentiate changes between AB and BB genotypes of the LGB gene, while for the CSN3 MMBRR had the best fit. Here, Kučerová et al. [24] concluded that the mixed model is the best suited for studies of genetic association with genes of casein; while such models shown less power to determine associations for genes of other milk-whey proteins, including β-lactoglobulin.

CONCLUSION

Differences in the magnitude of the influence of CSN3 and LGB genotypes depending on the breed could change according on the shared genetics. The mixed model based in Bayesian or Ridge Regression was the best alternative for analysis in genetic association studies involving the CSN3 and LGB genes. Due to their higher substitution allele effect and the minor inbreeding level, Brown Swiss and Fleckvieh could show more progress in selection programs that include the CSN3 and LGB genes, relative to Holstein or Jersey.

CONFLICT OF INTEREST

We certify that there is no conflict of interest regarding the material discussed in the manuscript.

ACKNOWLEDGMENTS

The authors thank Mexican Council of Science and Technology (CONACYT) for the financial support to the first author during his Doctoral studies.

REFERENCES

[1] Heck JML, Schennink A, Van Valenberg HJF, Bovenhuis H, Visker MHP, Van Arendonk JAM. Effects of milk protein variants on the protein composition of bovine milk. J Dairy Sci 2009;92:1192-202. DOI: 10.3168/jds.2008-1208

[2] Duifhuis-Rivera T, Lemus-Flores C, Ayala-Valdovinos MA, et al. Polymorphisms in beta and kappa casein are not associated with milk production in two highly technified populations of Holstein cattle in México. J Anim Plant Sci 2014;24:1316-21.

[3] Dogru U. β-lactoglobulin genetic variants in Brown Swiss dairy cattle and their association with milk yield and quality traits. J Anim Plant Sci 2015;25:595-98.

[4] Gustavsson F, Buitenhuis AJ, Johansson M, et al. Effects of breed and casein genetic variants on protein profile in milk from Swedish Red, Danish Holstein, and Danish Jersey cows. J Dairy Sci 2014;97:3866-77. DOI: 10.3168/jds.2013-7312

[5] Bernabucci U, Basiricò L, Morera P, et al. Effect of summer season on milk protein fractions in Holstein cows. J Dairy Sci 2015;98:1815–27. DOI: 10.3168/jds.2014-8788

[6] Streit M, Neugebauer N, Meuwissen THE, Bennewitz J. Evidence for a major gene by polygene interaction for milk production traits in German Holstein dairy cattle. J Dairy Sci 2011;94:1597-600. DOI: 10.3168/jds.2010-3834

[7] SAS. SAS 9.3 User´s Guide. Cary, NC: SAS Institute Inc; 2011.

[8] Poulsen NA, Buitenhuis AJ, Larsen LB. Phenotypic and genetic associations of milk traits with milk coagulation properties. J Dairy Sci 2015;98:1-9. DOI: 10.3168/jds.2014-7944

[9] Vidović V, Lukać D, Nemes Ž, Trivunović S. β-lactoglobulin genetic variants in Serbian Holstein Friesian dairy cattle and their association with yield and quality of milk. Anim Sci Pap Rep 2014; 32:179-82.

[10] Comin A, Cassandro M, Chessa S, et. al. Effects of composite β- and κ-casein genotypes on milk coagulation, quality, and yield traits in Italian Holstein cows. J Dairy Sci 2008;91:4022-27. DOI: 10.3168/jds.2007-0546

[11] Deb R, Singh U, Kumar S, Singh R, Sengar G, Sharma A. Genetic polymorphism and association of kappa-casein gene with milk production traits among Frieswal (HF x Sahiwal) cross breed of Indian origin. Iran J Vet Res 2014;15:406-8.

[12] Miciński J, Klupczyński J, Mordas W, Zablotna R. Yield and composition of milk from Jersey cows as dependent on the genetic variants of milk proteins. Pol J Food Nutr Sci 2007;57:95-9.

[13] Sitkowska B, Neja W, Milczewska A, Mroczkowski S, Markowska A. Milk protein polymorphisms and effect of herds on cow milk composition. J Central European Agr 2013;14:78-90.

[14] Raven L, Cocks BG, Hayes BJ. Multibreed genome wide association can improve precision of mapping causative variants underlying milk production in dairy cattle. BMC Genomics 2014;15:62. DOI: 10.1186/1471-2164-15-62

[15] Ramayo-Caldas Y, Renand G, Ballester M, Saintilan R, Rocha D. Multi-breed and multi-trait co-association analysis of meat tenderness and other meat quality traits in three French beef cattle breeds. Genet Sel Evol 2016;48:37.

[16] Chrenek P, Huba J, Vasicek D, Peskovicová D, Bulla J. The relation between genetic polymorphism markers and milk yield in Brown Swiss cattle imported to Slovakia. Asian-Australas J Anim Sci 2003;16:1397-401.

[17] Matějiček A, J Matějíčková, E Němcová, et. al. Joint effects of CSN3 and LGB genotypes and their relation to breeding values of milk production parameters in Czech Fleckvieh. Czech J Anim Sci 2007;52:83-7.

[18] Negrini R, Nijman IJ, Milanesi E, et. al. Differentiation of European cattle by AFLP fingerprinting. Anim Genet 2007;38:60-6. DOI: 10.1111/j.1365-2052.2007.01554.x

[19] Bonfatti V, Di Martino G, Cecchinato A, Vicario A, Carnier P. Effects of β-κ-casein (CSN2-CSN3) haplotypes and β-lactoglobulin (BLG) genotypes on milk composition traits and detailed protein composition of individual milk Simmental cows. J Dairy Sci 2010;93:3797-808. DOI: 10.3168/jds.2009-2778

[20] Oltenacu PA, Broom DM. The impact of genetic selection for increased milk yield on the welfare of dairy cows. Anim Welf 2010;19:39-49.

[21] Monir M, Zhu J. Comparing GWAS results of complex traits using full genetic model and additive models for revealing genetic architecture. Sci Rep 2017;7:38600. DOI: 10.1038/srep38600

[22] Oleński K, Cieślińska A, Suchocki T, Szyda J, Kamiński S. Polymorphism in coding and regulatory sequences of beta casein gene is associated with milk production traits in Holstein-Friesian cattle. Anim Sci Pap Rep 2012;30:5-12.

[23] Pärna E, Kaart T, Kiiman H, Bulitko T, Viinalass H. Milk protein genotype association with milk coagulation and quality traits. In: Chaiyabutr N editor. Milk Production-Advanced Genetic Traits, Cellular Mechanism, Animal Management and Health. Rijeka: InTech; 2012. p. 155-172.

[24] Kučerová J, Matějiček A, Jandurová OM, et. al. Milk protein genes CSN1S1, CSN2, CSN3, LGB and their relation to genetic values of milk production parameters in Czech Fleckvieh. Czech J Anim Sci 2006;51:241-7.

Table 1. Level of significance for the effects included in the models to analyze milk production (MP), and protein (PP) and fat (FP) percentage in milk

Variable	B[z]	M[y]	T[x]	AG[w]	GN[v]	SS[u]	BAG[t]	BGN[s]	MGN[r]
MP	0.1528	0.2411	<0.0001	0.1833	0.0974	<0.0001	<0.0001	<0.0001	0.0255
PP	0.7529	0.7358	0.6484	0.9966	1.000	0.9090			
FP	0.1512	0.4161	0.8409	0.8604	1.000	0.8199			

B[z], breed; M[y], model used in the original study; T[x], test used to recover original data; AG[w], analyzed gene; GN[v], genotype nested in analyzed gene; SS[u], sample size; BAG[t], interaction of breed by analyzed gene; BGN[s], interaction of breed by genotype nested in analyzed gene; MGN[r], interaction of model by genotype nested in analyzed gene.

Table 2. Least square means and multiple comparison for milk production for the interaction Breed by Gene

	Gen	
Breed	**CSN3**	**LGB**
Brown Swiss	5.60±0.46[ab]	5.64±0.46[ab]
Holstein	6.01±0.19[a]	6.02±0.19[a]
Jersey	5.55±0.25[b]	5.56±0.25[b]
Fleckvieh	5.98±0.22[ab]	5.70±0.22[ab]

[a-b] Means with different literal in the same row or column are different ($p<0.001$).

Table 3. Least square means, standard error, and multiple comparison for milk production for the subclasses of the interaction breed by genotype in the loci of the CSN3 gene and LGB gene

Breed	Genotype of the CSN3 gene			Genotype of the LGB gene		
	AA	**AB**	**BB**	**AA**	**AB**	**BB**
Brown Swiss	5.54±0.47[a]	5.62±0.47[a]	5.64±0.47[ab]	5.54±0.47[a]	5.62±0.47[ab]	5.64±0.47[ab]
Holstein	6.04±0.22[a]	6.01±0.22[a]	6.00±0.22[ab]	6.05±0.20[a]	6.01±0.20[a]	6.01±0.20[a]
Jersey	5.59±0.31[a]	5.57±0.31[a]	5.50±0.31[b]	5.60±0.19[a]	5.54±0.19[b]	5.57±0.19[b]
Fleckvieh	5.80±0.24[a]	5.86±0.24[a]	6.30±0.22[a]	5.71±0.22[a]	5.73±0.22[ab]	5.67±0.22[ab]

[a-b] Means with different letter in the same row or column are different ($p<0.001$).

Table 4. Least square means, standard error, and multiple comparison for milk production for the subclasses of the interaction model by genotype in the loci of the CSN3 gene and LGB gene

Model[1]	Genotype of the CSN3 gene			Genotype of the LGB gene		
	AA	AB	BB	AA	AB	BB
LS	5.96±0.18[a]	5.98±0.18[a]	6.09±0.18[a]	5.95±0.25[a]	5.98±0.25[a]	5.94±0.25[a]
MM	5.53±0.19[b]	5.54±0.19[b]	5.62±0.19[b]	5.54±0.19[a]	5.50±0.19[b]	5.50±0.19[a]

[1]LS, least square model; MM, mixed model based on bayesian and ridge regression.

[a-b] Means with different letter in the same row or column are different ($p<0.05$).

4. POTENTIAL CANDIDATE GENES FOR INTEGRAL GENETIC IMPROVEMENT PROGRAMS IN BEEF CATTLE

José Luis Zepeda-Batista[1], Gaspar Manuel Parra Bracamonte[2], Rafael Núñez Domínguez[1], Rodolfo Ramírez Valverde[1], Agustín Ruíz Flores[1]

[1]Universidad Autónoma Chapingo. Departamento de Zootecnia. Posgrado en Producción Animal. Km. 38.5 Carretera México-Texcoco. Chapingo, Estado de México. C.P. 56230

[2]Laboratorio de Biotecnología Animal. Centro de Biotecnología Genómica. Instituto Politécnico Nacional. Boulevard del Maestro S/N esq. Elías Piña, Col. Narciso Mendoza, Ciudad Reynosa, Tamaulipas, México, C.P. 88710.

Submitted to the *Italian Journal of Animal Science*

Potential candidate genes for integral genetic improvement programs in beef cattle

José Luis Zepeda Batista[1], Gaspar Manuel Parra Bracamonte[2], Rafael Núñez Domínguez[1€], Rodolfo Ramírez Valverde[1], Agustín Ruíz Flores[1]

[1]Universidad Autónoma Chapingo. Departamento de Zootecnia. Posgrado en Producción Animal. Km. 38.5 Carretera México-Texcoco. Chapingo, Estado de México. C.P. 56230

[2]Laboratorio de Biotecnología Animal. Centro de Biotecnología Genómica. Instituto Politécnico Nacional. Boulevard del Maestro S/N esq. Elías Piña, Col. Narciso Mendoza, Ciudad Reynosa, Tamaulipas, México, C.P. 88710.

[€] Correspondence author:

Rafael Núñez Domínguez. E-mail: rafael.nunez@correo.chapingo.mx

Abstract

Genetic association studies have successfully identified genetic variants associated with complex traits. Moreover, genome-wide association studies have allowed the identification of pleiotropic candidate genes using postGWAS procedures. Thus, this review intends to identify pleiotropic and consortium candidate genes associated with multiple economic importance traits and describe their potentially use in beef cattle breeding and integral genetic improvement programs. On this regard, less than 5% of the 400 genes reported in 81 genetic association studies were highly associated with one or more complex traits in beef cattle. Since pleiotropic and candidate genes consortiums have been reported to be associated with many of the innovative and functional traits that increase efficiency by reduced costs of input. Knowledge of their genetic architecture has made possible to figured out the genetic-biological mechanisms involved in such partnerships and choose the approach that best explains the phenomena. Additionally, pleiotropic and candidate genes consortiums have also shown high short term-benefits when they are used with traits that are not recorded on candidates for selection and thus their estimated breeding values for those traits have low accuracy. Therefore, the use of pleiotropic genes and candidate gene consortiums in integral genetic improvement programs could benefit commercial producers who needs short-term results.

Key words: consortium, multi-breed, multi-trait, pleiotropism

Introduction

In cattle, most of the economically important traits are genetically complex, which means they are polygenic. Polygenic traits are the basis of the genetic improvement and genomic assisted tools, several approaches to identify genetic polymorphisms associated with phenotypic differences have been proposed (1). Candidate gene and genome wide are the most common approaches used in genetic association studies (2). Recently, the approaches of candidate gene and genome wide has been combined allowing candidate gene studies to be the forefront of genetic association studies (2,3). Besides, candidate gene studies are relatively cheap and quick to perform and are focused on the selection of genes that have been in some way related to the trait previously and thus come with prior knowledge about gene function (3).

In beef cattle, candidate gene approach has been used to determine the effect of variations in the genome on traits such as birth, weaning and yearling weight (4), feeding efficiency (5), female and male reproduction traits (6), conformation (7), and carcass quality (8), among others. Due to the phenotypic variance explained by certain candidate genes and the relationship between feeding, growth, carcass quality, and other economical traits in beef cattle (9), this approach is an auxiliary tool in beef breeding programs for genetic improvement.

However, one of the main issues for integral genetic improvement is the observed genetic correlation among traits, resulting from the influence of polymorphisms affecting multiple complex traits (quantitative trait loci or QTL) (10,11). A positive genetic correlation that is less than 1.0 between two traits, such as weight and fatness, implies that some QTL affect both traits in the same direction, but other

QTL may affect only one trait and a small number may even affect the traits in the opposite direction (10). The response to selection reflects the genetic correlation between traits, which summarizes the genome-wide average effects of pleiotropy at shared loci (11). Identifying QTL with different patterns of pleiotropy should help us to understand the physiological control of multiple traits (10).

In this sense, the present study considered and analyzed 81 published articles in scientific journals between 2005 and 2017 that were related with GAS based on candidate genes previously or not identified by postGWAS procedures in beef cattle.

Thus, this review intended to identify pleiotropic and in consortium candidate genes (CG) associated with multiple economic importance traits (EIT) and described their potentially use in beef cattle breeding and integral genetic improvement programs.

Genetic association studies in beef cattle

In these last 20 years, genetic association studies (GAS) have been successfully used in identifying genetic variants associated with complex traits (12,13). Different approaches to GAS have been developed during this time, such as candidate gene, gene-wide, and genome-wide (1,2). Candidate gene and genome-wide approaches have been the most used in GAS. However, both approaches are based on different procedures to make the genetic association.

Candidate gene approach begins with the selection of a putative candidate gene based on a previous knowledge about gene function and its relevance in the mechanism of the studied trait. The first studies focused the attention on single

polymorphisms, in single genes, thought to have a major role, mainly located in exons, or in close regions regulating the gene expression (2,3)

On the other hand, the genome-wide approach uses sequence variations (mainly single nucleotide polymorphisms, SNPs) in the whole genome, together with the phenotype and pedigree information, to perform association analysis and to identify genes or regulatory elements that are important for the traits of interest. This approach provides the opportunity to cover the entire genome and to have information on genes involved in previously unsuspected pathways, which could have been never considered in a candidate gene approach (2,14).

However, there has been an increase of Genome-Wide Association Studies, that are using *in silico* and bioinformatic tools, such as the postGWAS procedures, to determine CG associated with EIT, reducing the biases of the original approach (3,13). These studies have allowed the identification of relationships between candidate genes that were not previously reported. Additionally, these studies could help in the identification of CG associated with more than one type of trait to use in genetic improvement programs.

Candidate genes associated with economically important traits in beef cattle

The results of the candidate genes studies in beef cattle have been variable, since they stablish associations in less than 1% of the analyzed markers in validation and identification studies for fatty acids in meat and up to 88% for carcass and growth traits (15,16), while in expression analysis with previously validated markers the associations were close to 100% (7,17). Phenotypic variance

reported as explained by the candidate genes known ranges from less than 1% to around 30%, depending on the analyzed trait (4,17,18).

Analyzed articles reported more than 400 genes, however only about 55% of them have been found to be informative or significantly associated with the phenotype in GAS using candidate genes in beef cattle. Besides, just 57 genes of those associated genes have shown similarities throughout the studies independently of the linkage disequilibrium and the breed present in the sample. Even so, more than a half of those genes explained very little of the phenotypic variance what made them less informative and applicable in GAS based on candidate genes. Nevertheless, the rest 40% of genes were highly associated with one or more complex traits (pleiotropic effect).

Since there are genes with pleiotropic effects (Tables 1 to 4) and aiming to simplify the analysis of those associations, genes could be associated with groups of traits that shared a common nature. In this sense, all traits of economic value should be considered when selecting beef cattle (19). On this regard, growth and conformation traits are usually considered as selection criteria in genetic improvement programs of beef cattle (20). Additionally, meat quality traits, like marbling and tenderness, are affected by factors such as animal feeding (21), and the management of carcasses during and after slaughter (22) conferring them great importance in genetic improvement programs.

Due to a common nature and their economic importance for breeder and beef industry, traits studied in the analyzed articles were grouped in 1) feeding traits (FT); 2) growth and conformation traits (GCT); 3) meat quality traits (MQT); and 4) reproductive traits (RT). Historically, most emphasis has been got on traits that

are most directly associated with profitability, and most easily measured, such as growth and conformation traits (23). However, recently, meat quality, feeding, and reproduction traits have been extensively studied due to their effect on the efficiency in the production and the possibility of maximizing the profits (24,25).

Regarding the present study, there were 10 CG associated with two groups of traits (Tables 1 and 2). However, there were three different subgroups of CG: those associated with CGT and RT, such as *CORIN* gene with the yearling weight in Mexican Charolais (26) and with the direct calving difficulty (DCD) in Irish Limousin, showing differences of 2.49% in the DCD PTA between animals with AA and CC genotype (27); genes associated with GCT and MQT, like *RORA* gene in 11 European cattle breeds, associated with the ratio of light absorption (K) to light scattering (S) at wavelengths between 670 and 740 nm, where greater S produces lower grades of meat quality. On this regard, animals with the AA genotype had greater scattering coefficients by 30 and 32% respectively, compared to AG, and GG genotypes (28).

Another gene influencing two groups of traits (MQT and FT) is *S100A10* gene, that in Chinese Simmental was mainly associated with fat and meat color, marbling score, *longissimus* muscle area, and shear force (29). Similarly, CG associated with three groups of traits (Table 3) showed sub-groups of association. *GHR* gene was associated with CGT, FT, and MQT in Nellore body weight gain, gross feed efficiency, residual feed intake, and carcass traits (30). Meanwhile, *PLXNB2* gene has been associated with birth and weaning weight (CGT), pre-weaning average daily gain (FT), pregnancy rate and calving difficulty (RT) in Angus-Charolais-Hereford crossbred cattle (6). Candidate genes associated with

CGT, FT, MQT, and RT (Table 4), like PLAG1 in Nellore cattle that explains from 0.29 to 2.53% of the bovine stature variability depending on the genotype present, did not have subgroups (31).

Candidate genes consortiums

Many genes have been used as a single or as groups of candidate genes in GAS studies. However, the use of GWAS tools has allowed the identification of multi-trait associated CG. It seems that some of those associations tend to depend on the genetic background of the studied animals (27). Besides, interaction network studies have determined the presence of pleiotropic or closely linked QTLs in multi-trait associated candidate genes (32). On this regard, there have been identified CG that behave as consortiums over several EIT (29). Additionally, those gene consortiums show interaction with other single CG like *LAP3* (33), *RPS20* (32), *LYN* (31), *GHR* (15,17), and *FABP4* (18), among others.

Twelve genes from the Tables 1 to 4, were identified to be part of a consortium or interact with one of them. Candidate gene consortiums identified and previously reported in the literature were Calpain (Calpain 1 or microcalpain) and Calpastatin (*CAPN1-CAST*), Calpain as the main regulator of postmortem proteolysis and Calpastatin as the regulator of Calpain (34); Coiled-coil-helix-coiled-coil-helix domain containing 7 and Pleiomorphic adenoma gene 1 (*CHCHD7-PLAG1*), both as important regulators in the pathway of growth and body height in cattle (35,36); Fatty acid binding protein 4 and Peroxisome proliferator-activated receptor gamma (*FABP4-PPARγ*), *FABP4* as a promotor of fatty acids deposition in the muscle and *PPARγ* as the regulator of *FABP4* (37); and Non-SMC condensin I

complex subunit G and ligand dependent nuclear receptor corepressor-like (*NCAPG-LCORL*), both as promotors of cattle growth, possibly through a role in cell proliferation expressed in liver, intestine and pancreas (24).

CAPN1-CAST

CAPN1 and *CAST* genes have been associated with several sensorial traits such as tenderness and juiciness in *Bos taurus* and *Bos indicus* cattle (34,38). However, it seems that both genes show a pleiotropic behavior and the highest association with other genes and traits. In this sense, *CAPN1* has shown high correlation with *ABCG2* with a significant effect on post-partum anestrus interval and ability ovulate prior to weaning in Brahman (39). Meanwhile, recently some studies determined that *CAPN1*, *CAST*, and *LEP* genes might be useful in marker assisted selection programs in Simmental cattle because of its high correlation between each other (38,40).

CHCHD7-PLAG1

The *CW-1* region in cattle has been reported as a region associated with carcass weight and stature (35,36). The *CHCHD7-PLAG1* genes have been targeted in resequencing analysis of that region concluding that both were the candidate causative genes of the reported association in the CW-1 region. Additionally, the pleiotropic behavior of those genes allowed them to influence CGT, FT, MQT, and RT in *Bos taurus* and *Bos indicus* breeds of cattle (36,41). Revealing in some cases the genetic architecture underlying of some complex traits or functioning as

a QTN explaining the phenotypic variance for those traits, *e.g.* until 8.6% of phenotypic variance for biceps in Chinese Simmental (35,36).

RPS20 is another CG in the *PLAG* region at 25 Mb on BTA 14 (4,31). It seems probable correlated with the *CHCHD7-PLAG1* consortium and similarly has shown a pleiotropic behavior both in European and Zebuine cattle (42). Besides, apparently its polymorphisms associated with CGT and RT explain until 10% of the paternal calving ease EBV variation in German Fleckvieh (42).

FABP4-PPARγ

Studies related to GAS of MQT and CGT have reported pleiotropic behavior and high influence of *FABP4-PPARγ* consortium (37,43). Due *FABP4* is involved in adipocyte differentiation in the *PPARγ* signaling pathway both are highly correlated between each other (43). Additionally, a gene co-expression network analysis made in Korean Hanwoo cattle determined the presence of high expressions of this gene-consortium in high-marbled cattle (43). Similarly, studies made in Aberdeen-Angus, Blonde D´Aquitaine, and Japanese Black cattle confirmed the influence of the *FABP4-PPARγ* consortium on fatty acids profile in the *Longissimus* muscle (37).

NCAPG-LCORL

The *NCAPG-LCORL* region is localized in the chromosome 6 and has been identified as a pleiotropic locus associated with CGT, MQT, and RT in breeds such as Piedmontese, Brangus, Simmental, and Japanese Black (16,33,44). On this regard, NCAPG-LCORL consortium showed a QTL at 38 Mb with large-

pleiotropic effect on growth traits in American Brangus population. With the identification of this QTL and its known biological pathways it was possible to determine the presence of biological processes for variation in GCT and MQT in Brangus cattle (16).

On the other hand, *NCAPG-LCORL* consortium has shown also high correlations with another CG, such as *LAP3* and *DCAF16*. In this sense, Italian Piedmontese cattle were analyzed and it was found the presence of high correlation between SNPs located in *LAP3* gene and the *NCAPG-LCORL* locus. Additionally, the *DCAF6-NCAPG* region was identified as a locus susceptible for average daily gain in Simmental cattle (45).

Genetic architecture and identification of consortium genes with pleiotropic effect

For years, genetic improvement programs were characterized by the selection of one or few traits considered as the main economic importance (46). However, this way of selective breeding can lead to mating difficulties, increasing of genetic disorders incidence and presence of dystocia in modern cattle (46). Additionally, the presence of negative correlations between some productive traits and functional traits illustrates the importance of knowledge about biological pathways and interaction networks between pleiotropic genes (47).

Recently, the genetic architecture of complex traits in cattle has been studied as an effort to obtained sufficient knowledge of gene function, determined pleiotropic gene effects on EIT, and determine possible QTN for those traits (48,49,50). Additionally, this knowledge has allowed understanding the effects of

hybridization between European and Indicine cattle on the EIT in commercial herds (51,52). The result of those studies is that genetic architectures of some genes with major effects are known in several cattle breeds (51).

During the last decade, genetic architecture studies have made possible to figured out the genetic-biological mechanisms of many production problems related with one- or two-trait selection programs (31,32,53). During this time low- and high-density genomic tools have been broadly used to study and characterize the genetic diversity and population structure of livestock (54). Besides, several strategies to evaluate and genetically improve cattle were developed (34,23,55). However, multi-trait, multi-breed, across-breed, and imputation analysis have been frequently used probing their high potential in GAS and integral genetic improvement programs in cattle.

Since the early 2010s-decade, multi-trait genomic analysis has been explored as a tool to identify the genetic correlations between complex traits measured in many breeding programs (8,56,57). On this regard, multi-traits GWAS was better than current single-traits genomic analysis selection breeding programs (8). Additionally, this approach has been able to identify pleiotropic patterns in genes (58) and increase the prediction accuracy of genomic values (8,59).

Multi-breed and across-breed analysis were other methods used to understand the genetic architecture of complex traits and identified pleiotropic patterns in genes. Besides, Bayesian-based multi-breed genomic models have better identified the correlation between traits and genes, increasing the accuracy of genomic prediction in distantly related breeds and small population breeds (60). Several studies showed the power of multi- and across-breed analysis to identify

genomic regions affecting GCT, FT, and MQT in purebred, crossbred, and composite breeds (61,62,63). On this regard, 173 core selective sweeps were identified in 37 breeds which may be used in crossbred or multi-breed genomic studies (49).

Recently, multi-breed and multi-trait co-association analysis were carried out in three French beef cattle breeds taking advantage of both strengths and the increasing of statistical power resolution (34). This approach identified 206 common candidate genes associated with GCT and MQT across the breeds confirming that genes associated with complex traits tend to be grouped together in clusters in the genome and present pleiotropic behavior (34,64).

Potential use of pleiotropic and candidate gene consortiums

Development of the SNP chip technologies and the identification and incorporation of Individual genes into selection schemes in livestock have increased our understanding of many economic importance species genetic architecture (65). However, the significant increased demand for animal source foods and considering that most livestock is produced by small holders, it is necessary the development of strategies to efficiently produced while avoiding genetic-based problems in production (54).

Pleiotropic and candidate genes consortiums have been reported to be associated with many of the innovative and functional traits like health, reproduction, meat quality, and behavior (53,66). Additionally, functional and innovative traits increase efficiency by reduced costs of input (66). On this regard, the economic benefit increased more than 100% when net merit values coming

from multiple-trait economic breeding objective analysis were used. In this sense, genomic selection indexes based on multi-trait and multi-breed (across-breed) co-analysis could substantially improve the present-day cost of genotyping the candidates for selection (65).

Depending on the specie and breed analyzed, multiple selection indexes are needed for different markets and production systems (67). Besides, due to the association of pleiotropic genes with multiple traits and the increasing in accuracy that can be obtained, the breeders could afford the investment of development a total performance index (TPI) (65,68). On this regard, pleiotropic and candidate genes consortiums have shown high short term-benefits when they are used with traits that are not recorded on candidates for selection and thus their estimated breeding values for those traits have a low accuracy (65).

Using correlated and uncorrelated phenotypes in the development and implementation of TPI let take advantage of major pleiotropic gene variants (50). Therefore, the use of pleiotropic and CG consortiums in integral genetic improvement programs could benefit commercial producers who need short-term results. On this regard, the MSTN gene could explain and improve up to 20% muscle mass trait (69), such gene could be used in crosses aimed to obtain animals with better conditions of development for commercial cattle herds, especially with breeds previously identified as carriers of mutations of MSTN gene, like Charolais cattle (26). Moreover, GHR gene influences FT in beef cattle obtaining improvements of 7 to 34% of commercial beef cattle herds, depending on the trait measured and the genotype of the animals (70).

The use of pleiotropic or consortium genes propitiate the analysis of more than one traits, then it is logical that commercial producer will also pay genomic services if they receive more benefits in a single test. The results between 9% to 76% of improvement obtained suggest an economic benefit when the cattle selection is based on a multi-trait breeding objective that considers pleiotropic and CG consortiums (65).

Conclusions and implications

Gene consortiums *CAPN1-CAST*, *CHCHD7-PLAG1*, *FABP4-PPARγ,* and *NCAPG-LCORL* could be used both as single or as a group in integral genetic improvement programs of economic importance traits due to their pleiotropic nature. Additionally, genes *LAP3, RPS20, LYN*, and *GHR*, interacts with candidate gene consortiums that could be used by small holders in low-cost breeding programs. In this sense, the pleiotropic and candidate gene consortiums proposed in this study could be used both for large or small herds that are genetically close or not.

Acknowledgments

The authors thank Mexican Council of Science and Technology (CONACYT) for the financial support to the first author during his Doctoral studies.

References

1. Paredes-Sánchez FA, Sifuentes-Rincón AM, Segura CA, García PCA, Parra BGM, Ambriz MP. Associations of SNPs located at candidate genes

to bovine growth traits, prioritized with an interaction networks construction approach. BMC Genet 2015;16:91.

2. Gianfagna F, Cugino D, Santimone I, Iacovello L. From candidate gene to genome-wide association studies in cardiovascular disease. Thromb Res 2012;129:320-324.
3. Patnala R, Clements J, Batra J. Candidate gene association studies: a comprehensive guide to useful *in silico* tools. BMC Genet 2013;14:39.
4. Fortes MRS, Reverter A, Hawken RJ, Bolormaa S, Lehnert SA. Candidate genes associated with testicular development, sperm quality, and hormone levels of inhibin, luteinizing hormone, and insulin-like growth factor 1 in Brahman bulls. Biol Reprod 2012;87(3):1-8.
5. Chen Y, Arthur PF, Herd RM, Quinn K, Barchia IM. Using genes differentially expressed in bulls to classify steers divergently selected for high and low residual feed intake. Anim Prod Sci 2012;52:608-612.
6. Akanno EC, Plastow G, Fitzsimmons C, Miller SP, Baron V, Ominski K, *et al*. Genome-wide association for heifer reproduction and calf performance traits in beef cattle. Genome 2015;58:549-557.
7. Maskur R, Arman C. Association of a novel single nucleotide polymorphism in growth hormone receptor gene with production traits in Bali cattle. Ital J Anim Sci 2015;13(4):3461.
8. Bolormaa S, Porto NLR, Zhang YD, Bunch RJ, Harrison BE, Goddard ME, *et al*. A genome-wide association study of meat and carcass traits in Australian cattle. J Anim Sci 2011;89:2297-2309.

9. Sharma A, Seop LJ, Gwon DC, Sudrajad P, Cheol KH, Heum YS, *et al*. Stories and challenges of genome wide association studies in livestock - A review. Asian Australas. J Anim Sci 2015;28(10):1371-1379.
10. Bolormaa S, Pryce JE, Reverter A, Zhang Y, Barendse W, Kemper K, *et al*. A multi-trait, meta-analysis for detecting pleiotropic polymorphisms for stature, fatness and reproduction in beef cattle. PLoS Genet 2014;10(3):e1004198.
11. Gratten J, Visscher PM. Genetic pleiotropy in complex traits and diseases: implications for genomic medicine. Genome Medicine 2016;8:78.
12. Visscher PM, Brown MA, McCarthy MI, Yang J. Five years of GWAS discovery. Am J Hum Genet 2012;90:7-24.
13. Pasaniuc B, Price AL. Dissecting the genetics of complex traits using summary association statistics. Nat Rev Genet 2017;18(2):117-127.
14. Zhang W, Li Y, Guo Y, Zhang L, Xu L, Gao X, *et al*. Multi-strategy genome-wide association studies identify the DCAF16-NCAPG region as a susceptibility locus for average daily gain in cattle. Sci Rep 2016;6:38073.
15. Karisa BK, Thomson J, Wang Z, Bruce HL, Plastow GS, Moore SS. Candidate genes and biological pathways associated with carcass quality traits in beef cattle. Can J Anim Sci 2013;93:295-306.
16. Weng Z, Su H, Saatchi M, Lee J, Thomas MG, Dunkelberger JR, *et al*. Genome-wide association study of growth and body composition traits in Brangus beef cattle. Livest Sci 2016;183:4-11.

17. Al-Husseini W, Gondro C, Quinn K, Herd RM, Gibson JP, Chen Y. Expression of candidate genes for residual feed intake in Angus cattle. Anim Genet 2013;45:12-19.
18. Maharani D, Jung Y, Jung WY, Jo C, Ryoo SH, Lee SH, *et al*. Association of five candidate genes with fatty acid composition in Korean cattle. Mol Biol Rep 2012;39:6113-6121.
19. Gadberry S, Jennings J, Ward H, Beck P, Kutz B, Troxel T.Beef cattle production. 1st ed. Arkansas, USA: The University of Arkansas System Division of Agriculture; 2016.
20. Barwick SA, Henzell AL. Development successes and issues for the future in deriving and applying selection indexes for beef breeding. Aust J Exp Agric 2005;45:923-933.
21. Guimaraes AL, Zerlotti MME, Carrilho CR, Branco RH, Pereira LML, dos Santos GJN, *et al*. Phenotypic association between feed efficiency and feeding behavior, growth and carcass traits in Senepol cattle. R Bras Zootec 2017;46(1):47-55.
22. Guerrero A, Velandia VM, Campo MM, Sañudo C. Some factors that affect ruminant meat quality: from the farm to the fork. Review. Acta Sci Anim Sci 2013;35(4):335-347.
23. Haskell MJ, Simm G, Turner SP. Genetic selection for temperament traits in dairy and beef cattle. Front Genet 2014; 5:368.
24. Lindholm-Perry AK, Sexten AK, Kuehn LA, Smith TPL, King DA, Shackelford SD, *et al*. Association, effects and validation of polymorphisms

within the *NCAPG* - *LCORL* locus located on BTA6 with feed intake, gain, meat and carcass traits in beef cattle. BMC Genet 2011;12:103.

25. Weber KL, Welly BT, Van Eenennaam AL, Young AE, Porto-Neto LR, Reverter A. *et al*. Identification of gene networks for residual feed intake in angus cattle using genomic prediction and RNA-seq. PLoS ONE 2016;11(3): e0152274.
26. Jahuey-Martínez FJ, Parra-Bracamonte GM, Sifuentes-Rincón AM, Martínez-González JC, Gondro C, García-Pérez CA, *et al*. Genomewide association analysis of growth traits in Charolais beef cattle. J Anim Sci 2016;94:4570-4582.
27. Purfield DC, Bradley DG, Evans RD, Kearney FJ, Berry DP. Genome-wide association study for calving performance using high-density genotypes in dairy and beef cattle. Genet Sel Evol 2015;47:47.
28. Sevane N, Armstrong E, Wiener P, Pong WR, Dunner S, GemQual Consortium. Polymorphisms in twelve candidate genes are associated with growth, muscle lipid profile and meat quality traits in eleven European cattle breeds. Mol Biol Rep 2014;41:4721-4731.
29. Xia J, Qi X, Wu Y, Zhu B, Xu L, Zhang L, *et al*. Genome-wide association study identifies loci and candidate genes for meat quality traits in Simmental beef cattle. Mamm Genome 2016;27(5-6):246-255.
30. Gomes RC, Silva SL, Carvalho ME, Rezende FM, Pinto LFB, Santana MHA, *et al*. Protein synthesis and degradation gene SNPs related to feed intake, feed efficiency, growth, and ultrasound carcass traits in Nellore cattle. Genet Mol Res 2013;12(3):2923-2936.

31. Utsunomiya YT, Carmo AS, Carvalheiro R, Neves HHR, Matos MC, Zavarez LB, *et al*. Genome-wide association study for birth weight in Nellore cattle points to previously described orthologous genes affecting human and bovine height. BMC Genet 2013;14:52.
32. Saatchi M, Schnabel RD, Taylor JF, Garrick DJ. Large-effect pleiotropic or closely linked QTL segregate within and across ten US cattle breeds. BMC Genomics 2014;15:442.
33. Bongiorni S, Mancini G, Chillemi G, Pariset L, Valentini A. Identification of a short region on Chromosome 6 affecting direct calving ease in Piedmontese cattle breed. PLoS ONE 2012;7(12):e50137.
34. Ramayo-Caldas Y, Renand G, Ballester M, Saintilan R, Rocha D. Multi-breed and multi-trait co-association analysis of meat tenderness and other meat quality traits in three French beef cattle breeds. Genet Sel Evol 2016;48:37.
35. Nishimura S, Watanabe T, Mizoshita K, Tatsuda K, Fujita T, Watanabe N, *et al*. Genome-wide association study identified three major QTL for carcass weight including the *PLAG1-CHCHD7* QTN for stature in Japanese Black cattle. BMC Genet 2012;13:40.
36. Song Y, Xu L, Chen Y, Zhang L, Gao H, Zhu B, *et al*. Genome-wide association study reveals the *PLAG1* gene for knuckle, biceps and shank weight in Simmental beef cattle. PLoS ONE 2016;11(12):e0168316.
37. Dujková R, Ranganathan Y, Dufek A, Macák J, Bezdícek J. Polymorphic effects of FABP4 and SCD genes on intramuscular fatty acid profiles in

longissimus muscle from two cattle breeds. Acta Vet Brno 2015;84:327-336.

38. Magalhaes AFB, Camargo GMF, Fernandes JGA, Gordo DGM, Tonussi RL, Costa RB, *et al*. Genome-wide association study of meat quality Traits in Nellore cattle. PLoS ONE 2016;11(6):e0157845.
39. Collis E, Fortes MRS, Zhang Y, Tier B, Schutt K, Barendse W, *et al*. Genetic variants affecting meat and milk production traits appear to have effects on reproduction traits in cattle. Anim Genet 2011;43(4):442-446..
40. Ardicli S, Dincel D, Samli H, Balci F. Effects of polymorphisms at LEP, CAST, CAPN1, GHR, FABP4 and DGAT1 genes on fattening performance and carcass traits in Simmental bulls. Arch Anim Breed 2017;60:61-70.
41. Fink T, Tiplady K, Lopdell T, JohnsonT, Snell RG, Spelman RJ, *et al*. Functional confirmation of PLAG1 as the candidate causative gene underlying major pleiotropic effects on body weight and milk characteristics. Sci Rep 2017;7:44793.
42. Pausch H, Flisikowski K, Jung S, Emmerling R, Edel C, Götz KU, *et al*. Genome-wide association study identifies two major loci affecting calving ease and growth-related traits in cattle. Genetics 2011;187:289-297.
43. Lim D, Chai HH, Lee SH, Cho YM, Choi JW, Kim NK. Gene expression patterns associated with peroxisome proliferator-activated receptor (*PPAR*) signaling in the *Longissimus dorsi* of Hanwoo (Korean cattle). Asian Australas J Anim Sci 2015;28(8):1075-1083.
44. Sasago N, Abe T, Sakuma H, Kojima T, Uemoto Y. Genome-wide association study for carcass traits, fatty acid composition, chemical

composition, sugar, and the effects of related candidate genes in Japanese Black cattle. Anim Sci J 2017;8:33-44.

45. Zhang H, Wang Z, Wang S, Li H. Progress of genome wide association study in domestic animals. J Anim Sci Biotechnol 2012;3:26.
46. Citek J, Hradecka E, Rehout V, Hanusova L. Obstetrical problems and stillbirth in beef cattle. Anim Sci Pap Rep 2011;2:109-118.
47. Barillet F. Genetic improvement for dairy production in sheep and goats. Small Rumin Res 2007;70:60-75.
48. De Koning DJ. The genetic architecture of economically important traits provides major challenges for the implementation of gene editing in livestock. Nat Inst Biosci J 2016;1:1-10.
49. Gutiérrez-Gil B, Arranz JJ, Wiener P. An interpretative review of selective sweep studies in *Bos taurus* cattle populations: identification of unique and shared selection signals across breeds. Front Genet 2015;6:167.
50. Xiang R, MacLeod IM, Bolormaa S, Goddard ME. Genome-wide comparative analyses of correlated and uncorrelated phenotypes identify major pleiotropic variants in dairy cattle. Sci Rep 2017;7:9248.
51. Porto-Neto LR, Reverter A, Prayaga KC, Chan EKF. Johnston DJ, Hawken RJ, *et al*. The genetics architecture of climatic adaptation of tropical cattle. PLoS ONE 2014;9(11):e113284.
52. Utsunomiya YT, Bomba L, Lucente G, Colli L, Negrini R, Lenstra JA, *et al*. Revisiting AFLP fingerprinting for an unbiased assessment of genetic structure and differentiation of taurine and zebu cattle. BMC Genet 2014;15:47.

53. Biscarini F, Nicolazzi EL, Stella A, Boettcher PJ, Gandini G. Challenges and opportunities in genetic improvement of local livestock breeds. Front Genet 2015;6(33):1-7.
54. Rothschild MF, Plastow GS. Applications of genomics to improve livestock in the developing world. Livest Sci 2014;166:76-83.
55. Van Raden PM, Tooker ME, Wright JR, Sun C, Hutchison JL. Comparison of single-trait to multi-trait national evaluations for yield, health, and fertility. J Dairy Sci 2014;97:7952-7962.
56. Calus MPL, Veerkamp RF. Accuracy of multi-trait genomic selection using different methods. Genet Selec Evol 2011;43:26.
57. Jia Y, Jannink JL. Multiple-trait genomic selection methods increase genetic value prediction accuracy. Genetics 2012;192:1513-1522.
58. Crispim AC, Kelly MJ, Guimaraes SEF, Silva FF, Fortes MRS, Wenceslau RR, *et al*. Multi-trait GWAS and new candidate genes annotation for growth curve parameters in Brahman cattle. PLoS ONE 2015;10(10):e0139906.
59. Pausch H, Emmerling R, Schwarzenbacher H, Fries R.Detailed phenotyping identifies genes with pleiotropic effects on body composition. BMC Genomics 2016;17:224.
60. Lund MS, Su G, Janss L, Guldbrandtsen B, Brondum RF. Invited review: Genomic evaluation of cattle in a multi-breed context. Livest Sci 2014;166:101-110.
61. Lu D, Sargolzaei M, Kelly M, Vander VG, Wang Z, Mandell I, *et al*. Genome-wide association analyses for carcass quality in crossbred beef cattle. BMC Genet 2013;14:80.

62. Lu D, Miller S, Sargolzaei M, Kelly M, Vander VG, Caldwell T, *et al*. Genome-wide association analyses for growth and feed efficiency traits in beef cattle. J Anim Sci 2013;91:3612-3633.
63. Veneroni-Gouveia G, Tizioto PC, Meirelles SLC, Santiago AC, Alencar MM, Regitano LCA. Candidate genes for carcass traits in a tropical-adapted Brazilian composite beef breed. Genet Mol Res 2015;14(4):16667-16674.
64. Wright D. The genetic architecture of domestication in animals. Bioinform Biol Insights 2015;9(S4):11-20.
65. MacNeil MD. Invited Review: Value of genomics in breeding objectives for beef cattle. R Bras Zootec 2016;45(12):794-801.
66. Kiplagat SK, Limo MK, Kosgey IS. 2012. Genetic improvement of livestock for milk production. In: Milk Production – Advanced Genetic Traits, Cellular Mechanism, Animal Management and Health. ed. INTECH. Ed. Kosgey *et al*. Rijeka, Croatia. p.78-96.
67. Van Raden PM. Invited review: selection on net merit to improve lifetime profit. J Dairy Sci 2004;87:3125-3131.
68. Safus P, Pribyl J, Veselá Z, Vostrý L, Stípková M, Stádník L. Selection indexes for bulls of beef cattle. Czech J Anim Sci 2006;51(7):285-298.
69. Sorbolini S, Bongiorni S, Cellesi M, Gaspa G, Dimauro C, Valentini A, *et al*. Genome wide association study on beef production traits in Marchigiana cattle breed. J Anim Breed Genet 2017;134:43-48.
70. Sherman EL, Nkrumah JD, Murdoch BM, Li C, Wang Z, Fu A, *et al*. Polymorphisms and haplotypes in the bovine neuropeptide Y, growth

hormone receptor, ghrelin, insulin-like growth factor 2, and uncoupling proteins 2 and 3 genes and their associations with measures of growth, performance, feed efficiency, and carcass merit in beef cattle. J Anim Sci 2008;86:1-6.

71. Lisa C, Albera A, Carnier P, Di Stasio L. Variability in candidate genes revealed associations with meat traits in the Piedmontese cattle breed. Italian J Anim Sci 2013;12(2):e46.

72. Cesar ASM, Regitano LCA, Mourao GB, Tullio RR, Lanna DPD, Nassu RT, *et al*. Genome-wide association study for intramuscular fat deposition and composition in Nellore cattle. BMC Genet 2014;15:39.

Tables

Table 1. Candidate genes associated with two groups of traits usually considered in genetic improvement programs

Symbol	Name	Chr	Location	Associated Traits	References
CAPN1	Calpain 1	29	*Bos Taurus:* AC_000186.1 (44063463...44100316)	BT, BWT, Crc, InF, LMD, MLP, Mrb, REA, Tnd, WW, YW	8,28,61
CAST	Calpastatin	7	*Bos Taurus:* AC_000164.1 (98444826..98581260) *Bos indicus*: NC_032656.1 (97310302..97446918)	Crc, BWT, InF, MLP, Mrb, Tnd, WW, YW	8,28
CORIN	Corin, serine peptidase	6	*Bos Taurus:* AC_000163.1 (67923558..68232043) *Bos indicus:* NC_032655.1 (68875619..69192702)	BW, FRT, WW, YW	15,27
CRH	Corticotropin releasing hormone	14	*Bos Taurus:* AC_000171.1 (32213146..32609871)	BT, BWT, Conf, MLP, REA, WW, YW	28,63
FABP4	Fatty acid binding protein 4, adipocyte	14	*Bos Taurus:* AC_000171.1 (46833665..46838053)	BT, Conf, FAM, Mrb, REA	18,63

BT: backfat thickness; BWT: birth weight; Crc: Conf: conformation; FRT: female reproductive traits carcass; InF: intramuscular fat; LMD: longissimus muscle development; MLP: muscle lipid profile; Mrb: marbling; FAM: fatty acids in meat; REA: rib eye area; Tnd: tenderness; WW: weaning weight; YW: yearling weight.

Table 2. Candidate genes associated with two groups of traits usually considered in genetic improvement programs (II)

Symbol	Name	Chr	Location	Associated Traits	References
LAP3	Leucine aminopeptidase 3	6	*Bos Taurus:* AC_000163.1 (38574590..38600027) *Bos indicus*: NC_032655.1 (37963313..37988779)	BW, FAM, WW, YW	33,36
LEP	Leptin	4	AC_000161.1 (93249803..93266625)	BT, Crc, REA	63,71
MSTN *(GDF8)*	Myostatin	2	*Bos taurus:* AC_000159.1 (6213566..6220196) *Bos indicus*: NC_032651.1 (6524051..6530697)	BW, MLP, WW, YW	32,69
RORA	RAR related orphan receptor A	10	*Bos* taurus: AC_000167.1 (48949517..49762212) *Bos* indicus: NC_032659.1 (49971191..50076652)	BW, FAM, MLP, WW, YW	28,72
S100A10	S100 calcium binding protein A10	3	*Bos* taurus: AC_000160.1 (18799612..18810545) *Bos* indicus: NC_032652.1 (20271568..20283090)	FAM, REA, RFI, Mrb, Tnd	5,17,29

BT: backfat thickness; BWT: birth weight; Crc: carcass; FAM: fatty acids in meat; MLP: muscle lipid profile; Mrb: marbling; REA: rib eye area; RFI: residual feed intake; Tnd: tenderness; WW: weaning weight; YW: yearling weight.

Table 3. Candidate genes associated with three groups of traits usually considered in genetic improvement programs

Symbol	Name	Chr	Location	Traits	References
ABCG2	ATP binding cassette subfamily G member 2	6	Bos taurus: AC_000163.1 (37902882..38030585) Bos indicus: NC_032655.1 (37296106..37423294)	ADG, BT, BWT, Crc, InF, LMD, MLP, Mrb, REA, Tnd, WW	5,16,17,24, 29,32,35
GHR	growth hormone receptor	20	AC_000177.1 (31890736..32064204)	ADG, BWT, Crc, Conf, FAM, InF, Mrb, REA, Tnd, WW, YW	7,15,17, 44,71
LCORL	ligand dependent nuclear receptor corepressor like	6	AC_000163.1 (38840864..38992112)	BW, FAM, MLP, WW, YW	14,16,24, 32,33,35
NCAPG	non-SMC condensin I complex subunit G	6	AC_000163.1 (38711560..38812056)	BW, Crc, Conf, FAM, MLP, REA, WW, YW	14,16,35
PLXNB2	plexin B2	5	AC_000162.1 (119840726..119863921)	ADG, BWT, FRT, WW	6,27

ADG: average daily gain; BT: backfat thickness; BWT: birth weight; Crc: carcass; Conf: conformation; FAM: fatty acids in meat; FRT: reproductive traits; InF: intramuscular fat; LMD: longissimus muscle development; MLP: muscle lipid profile; Mrb: marbling; REA: rib eye area; Tnd: tenderness; WW: weaning weight; YW: yearling weight.

Table 4. Candidate genes associated with four groups of traits usually considered in genetic improvement programs

Symbol	Name	Chr	Location	Traits	References
CHCHD7	coiled-coil-helix-coiled-coil-helix domain containing 7	14	*Bos Taurus:* AC_000171.1 (25052830..25058781*)* *Bos indicus*: NC_032663.1 (23329319..23336283)	ADG, BWT, Crc, FRT WW, YW	31,35
LYN	LYN proto-oncogene, Src family tyrosine kinase	14	AC_000171.1 (24847257..24921758)	ADG, BWT, FAM, FRT, WW, YW	31,42
PLAG1	PLAG1 zinc finger	14	AC_000171.1 (25000459..25052403)	ADG, BT, BWT, Crc, FAM, FRT, LMD, REA, RFI, TD, WW, YW	4,14,31,32,35, 36,42,44,52
RPS20	ribosomal protein S20	14	AC_000171.1 (24955076..24956324)	ADG, BWT, Crc, FRT, TD, WW, YW	4,31,32,42

ADG: average daily gain; BT: backfat thickness; BWT: birth weight; Crc: carcass; FAM: fatty acids in meat; FRT: reproductive traits; LMD: longissimus muscle development; REA: rib eye area; RFI: residual feed intake; TD: testicular development; WW: weaning weight; YW: yearling weight.

5. SCREENING OF GENETIC DISEASES PREVALENCE IN BRAUNVIEH CATTLE

J.L. Zepeda-Batista[1], G.M. Parra-Bracamonte[2], R. Núñez-Domínguez[1], R. Ramírez-Valverde[1], A. Ruíz-Flores[1]

[1]Universidad Autónoma Chapingo. Departamento de Zootecnia. Posgrado en Producción Animal. Km. 38.5 Carretera México-Texcoco. Chapingo, Estado de México. C.P. 56230

[2]Laboratorio de Biotecnología Animal. Centro de Biotecnología Genómica. Instituto Politécnico Nacional. Boulevard del Maestro S/N esq. Elías Piña, Col. Narciso Mendoza, Ciudad Reynosa, Tamaulipas, México, C.P. 88710.

Submitted to the *Tropical Animal Health and Production*

Screening genetic diseases prevalence in Braunvieh cattle

Genetic diseases prevalence in Braunvieh cattle

José Luis Zepeda-Batista[1]: jzepedab@gmail.com
Gaspar Manuel Parra-Bracamontes[2]: gparra@ipn.mx
Rafael Núñez-Domínguez[1]: rafael.nunez@correo.chapingo.mx
Rodolfo Ramírez-Valverde[1]: rodolforv@correo.chapingo.mx
Agustín Ruíz-Flores[1]: arf@correo.chapingo.mx

[1]Universidad Autónoma Chapingo. Departamento de Zootecnia. Posgrado en Producción Animal. Km. 38.5 Carretera México-Texcoco. Chapingo, Estado de México. C.P. 56230
[2]Centro de Biotecnología Genómica. Instituto Politécnico Nacional. Boulevard del Maestro S/N esq. Elías Piña, Col. Narciso Mendoza, Ciudad Reynosa, Tamaulipas, México, C.P. 88710.

Corresponding author:

Rafael Núñez-Domínguez[1]: rafael.nunez@correo.chapingo.mx

Abstract

Background: Heritable abnormalities can cause a reduction in productive performance, structural defects or death of the animal. There are reports of hereditary abnormalities in bovine European Brown Swiss from several countries, but no evidence was found on their existence in Mexico. Identification and elimination of affected animals are important issues in genetic improvement programs. However, carrier animals are the main obstacle for the breeders to control the dissemination or achieve the elimination of the disease in the population.

Results: A total of 28 genes associated with hereditary diseases were screened with the GGP-LD 30K array (GeneSeek®) in 300 Mexican registered Braunvieh animals. Allelic frequencies of the markers associated with illness were obtained for: citrullinaemia, spinal dysmyelination, spinal muscular atrophy, Brows Swiss fertility haplotype 2, congenital muscular dystonia, epidermolysis bullosa, Pompes, maple syrup urine, syndactyly, Weaver syndrome, crooked tail, deficiency of uridine monophosphate synthase, hypotrichosis, Marfan syndrome,

and weak calf syndrome. The allelic frequency values were low for all the analysed loci (from 0.0015 to 0.0110), with exception of syndactyly (0.4145). Although homozygous animals for these genetic conditions were detected, no physical or physiological abnormalities associated with the clinical form of the diseases were observed in the sampled animals. Markers associated with crooked tail, deficiency of uridine monophosphate synthase, hypotrichosis, Marfan syndrome, and weak calf syndrome were absent.

Conclusions: The studied Mexican Braunvieh population does not present clinical or subclinical effects for 10 diseases in homozygous animals. However, since the assessed animals are considered as breeding stock, the monitoring of carrier animals might be periodically necessary.

Key words: Molecular Diagnosis, SNPs, Spinal Dysmyelination, Syndactyly, Weaver Syndrome

1. Introduction

The cattle breed known internationally as Braunvieh has its origins in the Alps region of several Central-Europe countries. In Mexico, original Braunvieh (OBV) cattle were introduced from Switzerland in the mid-nineteen hundred. On this regard, the *Asociación Mexicana de Criadores de Ganado Suizo de Registro* (AMCGSR) (1) reported that Braunvieh is one of the cattle breeds most used in the Mexican beef production industry, either as purebred or in crosses with *Bos indicus* cattle. However, despite the long time that Braunvieh cattle has been used in Mexico, there is scarce information on the breed productive performance, and the available information is mostly related with growth traits coming from genetic evaluations (1). Thus, none information is available regarding genetic diseases prevalence in the Mexican Braunvieh cattle population.

Genetic abnormalities contribute to poor animal performance, structural unsoundness, and semi-lethal or lethal diseases (2,3). Identification and elimination of affected animals are important issues in genetic improvement

programs. However, carrier animals are the main obstacle for the breeders to control the dissemination or achieve the elimination of the disease in the population (4).

Braunvieh cattle is one of the breeds affected with genetic diseases, especially in Europe and USA (5). On this regard, European and American populations of Braunvieh cattle are a mixture of OBV and Brown Swiss (BSW) (6). Some authors (7,8,9) mentioned that BSW genetic flow in the OBV population causes the incidence of some genetic diseases which could not be found previously in the American and European Braunvieh populations. Cole *et al*. (10) reported almost USD$11 million of economic losses per year in the USA caused by reduced fertility and affected embryos with genetic diseases. Then, despite the OBV genetic base of the Mexican population, the common use of artificial insemination and the genetic flow from Austria, Swiss and US, make it necessary to determine the status of the genetic disease of the Mexican Braunvieh cattle.

The objectives of the present study were to determine the existence, and the allelic and genotypic frequencies of 28 genetic diseases of economic importance in the Mexican Braunvieh cattle population.

2. Material and methods

2.1 Samples for DNA extraction

Hair follicles samples were collected from 300 Braunvieh individuals registered in the AMCGSR. Genetic background of the sampled population included Austrian, Swiss, Canadian, American and Mexican animals. The cattle were born between 2001 and 2016, and included 236 females and 64 males from five herds located in Eastern, Central and Western Mexico. Sampled cows had at least a calf, whereas, sires had at least two calves in different herds.

2.2 Genotyping of animals

The animals were genotyped using 30,125 SNP markers from the GeneSeek® Genomic Profiler Bovine LD v4 panel (Neogen Corp. Lincoln, NE, USA). Twenty-eight genetic diseases previously reported for Braunvieh and other beef cattle breeds were considered in the study (Table 1). Sixty-three markers included in the SNP array and previously associated with the studied diseases were used to perform the screening.

2.3 Data analysis

Allelic and genotypic frequencies were estimated using the software CERVUS 3.0.7 (11).

3. Results

The results of the 28 genetic diseases screened in the Mexican Braunvieh population showed the presence of markers positively associated to 15 of them. Markers previously associated with genetic diseases present in the studied population, and their genetic and allelic frequencies are shown in Table 2. Genetic diseases with no presence of associated heterozygotes or homozygotes in the sample studied were Angus dwarfism, α- and β-mannosidosis, bovine arachnomelia syndrome, Chediak-Higashi syndrome, congenital myasthenic syndrome, dilated cardiomyopathy, factor XI, glycogen storage disease V, protoporphyria, pseudomyotonia, bovine leukocyte adhesion deficiency, and chondrodysplastic dwarfism.

Bovine citrullinaemia (Cit), bovine spinal dysmyelination (BSD), bovine spinal muscular atrophy (SMA), congenital muscular dystonia (CMD), crooked tail syndrome (CTS), deficiency of uridine monophosphate synthase (DUMPS), epidermolysis bullosa (EB), hereditary perinatal weak calf syndrome (WCS), hypotrichosis (Hyp), bovine Marfan syndrome (BMS), maple syrup urine disease (MSUD), and glycogen storage disease II (Pompes), were only identified in animals from west and central west Mexico herds. While Braunvieh fertility

haplotype (BH2) was identified just in the eastern Mexico herds. Meanwhile, markers of bovine progressive degenerative myeloencephalopathy (WS), and syndactyly syndrome (MF), were widespread throughout the population.

4. Discussion

Genetic diseases occur in all breeds of cattle (3). Differences in the prevalence of the disease of the herds in the population studied could result from the source of the germplasm used in the improvement programs in each region. Most of the semen and sires used in the east Mexico herds came from the southeast Mexican herds and just one herd from central highlands of Mexico. On the contrary, western and central western herds, generally used Mexican sires from other herds of the same region or foreign sires, especially from the U.S. and Canada.

Epidermolysis bullosa is one of the diseases with presence reported in both beef and dairy cattle. Laminin subunit 2 (*LAMC2*; GenBank accession no. **AC_000173.1**) gene has been one of the genes reported in beef cattle associated with EB. On this regard, Murgiano *et al.* (12) reported injuries associated with EB in homozygous animals for *LAMC2* gene. In the present study there was one homozygous animal for EB, 0.3%, but it did not show signs of the disease. The absence of injuries in homozygous animals could be associated with the marker that was positive in this study, “EB_2”. Since the marker “EB_3” was reported to cause the clinical disease in homozygous animals of breeds like Chianina and Hereford (12).

Pompes was another disease for which one homozygous animal was identified, 0.3%. Citek *et al*. (13) reported an absence of Pompes affected animals in seven Czech beef cattle breeds. However, the homozygous animal in the present study did not seem to be affected either by the clinical or subclinical form of the disease. This peculiarity matches with the published by Brooks and Koeberl (14), who mentioned that Pompes might be breeds-specific; thus, not all breeds would develop the disease.

Braunvieh fertility haplotype 2 associated markers were also identified in the studied population. Cole *et al*. (10) found no phenotypic or genetic effects of carriers for BH1 and BH2 haplotypes in Brown Swiss cattle. A similar behaviour was shown by the carriers and non-carriers of associated markers to CMD in the studied population. On this regard, Drögemüller *et al.* (15) reported the absence of CMD in the Brown Swiss cattle they studied and there were not recognizable effects on the animals.

In the present study it was determined the existence of heterozygous and homozygous animals for markers associated with MSUD. Several authors (16,17,18) reported the presence MSUD of affected animals in American, Canadian, Australian, Argentinian and Uruguayan populations of Polled Hereford, Polled Shorthorn and their crosses. However, these authors reported a disease incidence frequency between 1 to 2%, whereas in the present study the frequency was 0.3% for MSUD and MSUD_3 heterozygote and MSUD_2 homozygote markers. In addition, homozygous animals for the MSUD_2 in the present study did not show clinical signs of those illnesses. A possible explanation could be, as in other diseases, that not all associated disease markers lead to a clinical form regardless of the breed.

Bovine Marfan syndrome was another disease present in the studied population. Hirano *et al*. (19) reported 84.7% of carriers and 14.9% of affected homozygous animals in a Japanese Wagyu population. On the contrary, the frequency of BMS in the Braunvieh studied population was quite low, 0.3%. Likewise, it was identified the presence of associated markers with Hyp. Some authors reported affected animals in crossbred cattle, especially in animals Red Angus-Charolais-Simmental (20) and Hereford-Friesian (21) crossbred animals. In this study, only one heterozygous animal for Hyp was found, 0.3%. On this regard, given that Hyp expresses itself as complete or partial loss of hair, some breeders assume their animals carry the slick coat gene.

Another disease that probably has been indirectly selected for is CTS. Sartelet *et al*. (22) reported that markers associated with CTS had favourable effects on muscularity in heterozygotes. All the animals that showed the presence

of CTS-associated markers were heterozygotes, 0.14%, and they were susceptible to selection because of their outstanding conformation. Then, regarding the Hyp and CTS loci, it is possible that most of the carrier animals had been selected as replacements.

Some diseases, such as Cit and DUMPS are mainly associated with dairy cattle. However, Meydan *et al*. (23) reported the absence of carriers for both diseases in Turkish Holstein. On the contrary, in the present study there were homozygous animals for Cit and heterozygous animals for DUMPS, both diseases with allelic frequencies of 0.3%. A peculiarity of Cit was that the homozygous animal in the studied population did not show clinical signs associated with the disease. Uffo *et al*. (24) mentioned that Cit and DUMPS were Holstein-specific, meaning that no signs and symptoms are shown in animals without Holstein genetics.

After the screening, the result of this study allowed to identify the presence of positive animals for associated markers with BSD, SMA, WCS and WS. Those diseases are important for the studied population because of their strong association with the Brown Swiss Cattle, in general (2). It is necessary to keep in mind that BVH cattle in Mexico has been crossed with other Bos indicus and Bos taurus breeds for several years in absorption crossbreeding, especially with BSW, Brahman, and Nelore (1). Resulting progeny have reached the breed composition to be registered as purebred animals; however, genes of the other breeds are now integrated into the germplasm of Mexican BVH. In this sense, BSD, SMA, and WS are inherited of central nervous diseases usually reported in Braunvieh or crossbred calves upgraded with BSW (4).

Nissen *et al*. (25) reported the clinical form of BSD in homozygous BVH x BSW calves in Germany and the need to identify carrier animals to exclude from the breeding programs. Similarly, Krebs *et al*. (26) reported affected calves with SMA in crossbred BVH x BSW herds from the US and Europe. Observed frequencies in the present study of both markers diseases, BSD and SMA, were low (0.3%). Thomsen *et al*. (9) indicated that only the homozygous animals showed the clinical form of the BSD. Meanwhile, Medugorac *et al*. (27) reported

the same pattern for SMA. An exploratory pedigree analysis of the sample studied indicated that those animals positively associated with BSD and SMA only have OBV ancestors and none of them were upgraded with BSW cattle at least for the last four generations. However, none of the animals in this study showed the clinical form of the diseases.

Weak calf syndrome and WS are other diseases reported as breed-associated in purebred and cattle upgraded to BSW (2,4). Results of the present screening indicated the presence of heterozygous animals for markers associated with both diseases. The frequency for WCS in this study was 0.3%, whereas Hirano *et al*. (28) determined 7% of carrier animals in a Japanese Black cattle population. Frequencies obtained for WCS showed that population is in risk, especially for the disease ability not only to cause calf death but also embryonic or foetal death (29). On the other hand, WS in the studied population showed a frequency of 0.8% of heterozygous animals. This frequency was higher than the reported by Kunz *et al*. (8) who estimated a 0.26% of carrier animals in purebred German BSW. Nevertheless, in the present study 0.7% of the studied animals were homozygous for markers WVR_49657798, WVR_49691015, WVR_49692485, and WVR_49695952. Even so, none of them showed evident physical or physiological abnormalities associated with the clinical form of the disease.

Finally, the present screening also identified MF associated markers, showing the highest frequencies through the analysed population. Previously, MF has been associated with the LDL receptor related protein 4 gene (*LRP4*; GenBank accession no. **AC_000172.1**) and reported in both beef and dairy cattle breeds (30,31). Drögemüller et al. (31) reported four families of mutations in *LRP4* gene (Holstein I, Holstein II, Simmental, and Crossbred families), each of which had different segregation and limb-damage development level.

In the present study, animals with mutations in exons 3 (c.241G>A) and 20 (c.2719G>A) of the *LRP4* gene for MF were present. Homozygous and heterozygous individuals did not present apparent physical abnormalities. Drögemüller *et al*. (31) considered changing c.241G>A (M241) as a benign

mutation with no clinical signs of the MF. In addition, the disease variant of the M241 marker has the major incidence in the studied population. In this sense, MF might have been under indirect selection due to its previously reported favourable association with growth traits (32,33). Additionally, Drögemüller *et al*. (31) consider c.2719G>A (M2719) as a possibly damaging mutation, mainly in crossbred animals. However, M2719 associated markers in the present study were shown only in heterozygotes animals. Some authors (31,33) reported that MF mutations are frequent in an inbred population. Ruíz-Flores *et al*. (34) reported in the Mexican Braunvieh population, 3.1 ± 5.0% of inbreeding, thus rising inbreeding could set the conditions of homozygous animals emergence for the disease variant of the MF marker M2719.

5. Conclusion

Homozygous animals for Bovine Spinal Dysmyelination, Braunvieh Fertility Haplotype 2, Citrullinemia, Congenital Muscular Dystonia, Epidermolysis Bullosa, Mule Foot Syndrome, Maple Syrup Urine Disease, Spinal Muscular Atrophy, Pompes, and Bovine Progressive Degenerative Myeloencephalopathy were present in the studied population. Braunvieh cattle did not develop clinical or subclinical forms, even if animals were upgraded with BSW and Brahman cattle. Frequencies obtained in the studied population suggest implementing tracking of animals for the sake of avoiding possible spread of the associated disease genes. Since these animals are breeding stock, depending on the frequency of reproduction and sex, the spreading of genetic conditions needs to be periodically monitored, otherwise segregation would be unnoticed through the herd or population.

Additionally, it is necessary to implement a disease test in the Braunvieh population to avoid the indirect selection of disease genes, particularly for Hypotrichosis, Crooked Tail Syndrome and Syndactyly Syndrome. Results show a possibility to implement a scheme of molecular assisted diagnosis in the population.

6. Financial support

This work was supported by Universidad Autónoma Chapingo, Mexico (DGIP-166701012). The authors also thank Mexican Council of Science and Technology (CONACYT) for the financial support to the first author during his Doctoral studies and the Asociación Mexicana de Criadores de Ganado Suizo de Registro.for allowing the use of their databases and the collaboration of the breeders in the present study.

7. References

[1] AMCGSR (Asociación Mexicana de Criadores de Ganado Suizo de Registro). (http://www.amcgsr.com.mx/del-suizo-europeo.php). Consulted on October 28th, 2017.

[2] Gholap PN, Kale DS, Sirothia A. Genetic diseases in cattle: a review. Res J Anim Vet Fish and Sci 2014;2:24–33.

[3] Ciepłoch A, Rutkowska K, Oprządek J, Poławska E. Genetic disorders in beef cattle: A review. Genes Genomics 2017;39:461–71. doi:10.1007/s13258-017-0525-8.

[4] Gentile A, Testoni S. Inherited disorders of cattle: A selected review. Slov Vet Res 2006;43:17-29.

[5] Schwarzenbacher H, Burgstaller J, Seefried FR, Wurmser C, Hilbe M, Jung S, et al. A missense mutation in *TUBD1* is associated with high juvenile mortality in Braunvieh and Fleckvieh cattle. BMC Genomics 2016;17:400. doi:10.1186/s12864-016-2742-y.

[6] Maxa J, Neuditschko M, Russ I, Förster M, Medugorac I. Genome-wide association mapping of milk production traits in Braunvieh cattle. J Dairy Sci 2012;95:5357–64. doi:10.3168/jds.2011-4673.

[7] Drögemüller C, Tetens J, Sigurdsson S, Gentile A, Testoni S, Lindblad-Toh K, et al. Identification of the bovine arachnomelia mutation by massively parallel sequencing implicates sulfite oxidase (SUOX) in bone development. PLoS Genet 2010;6:6–12. doi:10.1371/journal.pgen.1001079.

[8] Kunz E, Rothammer S, Pausch H, Schwarzenbacher H, Seefried FR, Matiasek K, et al. Confirmation of a non-synonymous SNP in *PNPLA8* as a candidate causal mutation for Weaver syndrome in Brown Swiss cattle. Genet Sel Evol 2016;48:21. doi:10.1186/s12711-016-0201-5.

[9] Thomsen B, Nissen PH, Agerholm JS, Bendixen C. Congenital bovine spinal dysmyelination is caused by a missense mutation in the *SPAST* gene. Neurogenetics 2010;11:175–83. doi:10.1007/s10048-009-0214-0.

[10] Cole JB, Null DJ, VanRaden PM. Phenotypic and genetic effects of recessive haplotypes on yield, longevity, and fertility. J Dairy Sci 2016;99:7274–88. doi:10.3168/jds.2015-10777.

[11] Kallinowski ST, Taper ML, Marshall TC. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Mol Ecol 2007;16:1099-1106. doi: 10.1111/j.1365-294x.2007.03089.x.

[12] Murgiano L, Wiedemar N, Jagannathan V, Isling LK, Drogemuller C, Agerholm JS. Epidermolysa bullosa in Danish Hereford calves is caused by a deletion in *LAMC2* gene. BMC Vet Res 2015;11:334. doi:10.1186/s12917-015-0334-8.

[13] Citek J, Rehout V, Vecerek L, Hajkova J. Genotyping glycogen storage disease type II and type V in cattle reared in the Czech Republic. J Vet Med A Physiol Pathol Clin Med 2007;54:257–9. doi:10.1111/j.1439-0442.2007.00931.x.

[14] Brooks ED, Koeberl DD. Large animal models and new therapies for glycogen storage diseases. J Inherit Metab Dis 2015;38(3):505-509. doi: 10.1007/s10545-014-9766-8.

[15] Drögemüller C, Drögemüller M, Leeb T, Mascarello F, Testoni S, Rossi M, et al. Identification of a missense mutation in the bovine *ATP2A1* gene in congenital pseudomyotonia of Chianina cattle: An animal model of human Brody disease. Genomics 2008;92:474–7. doi:10.1016/j.ygeno.2008.07.014.

[16] Healy PJ, Dennis JA, Windsor PA, Pierce KD, Schofield PA. Genotyping cattle for inherited congenital myoclonus and maple syrup urine disease. Aust Vet J 2002;80:695–7.

[17] Dutra F, Romero A, Quinteros C, Kelly L. MSUD (Maple Syrup Urine Disease) en terneros Polled Hereford y cruzas Polled Hereford x Shorthorn en Uruguay. SMvU 2015;51:14–25.

[18] Robarge, M. E., J. E. Beever, S. D. Lenz, C. J. Lynch, and W. Wigle. 2015. Maple syrup urine disease in a central Indiana Hereford herd. Case Rep Vet Med 204037: 1-4. doi: 10.1155/2015/204037

[19] Hirano T, Matsuhashi T, Kobayashi N, Watanabe T, Sugimoto Y. Identification of an *FBN1* mutation in bovine Marfan syndrome-like disease. Anim Genet 2011;43:11-17. doi: 10.1111/j.1365-2052.2011.02209.x.

[20] Barlund CS, Clark EG, Leeb T, Drögemüller C, Palmer CW. Congenital hypotrichosis and partial anodontia in a crossbred beef calf. Can Vet J 2007;48:612–4.

[21] Jolly RD, Wills JL, Kenny JE, Cahill JI, Howe L. Coat-colour dilution and hypotrichosis in Hereford crossbred calves. N Z Vet J 2008;56:74–7. doi:10.1080/00480169.2008.36812.

[22] Sartelet A, Klingbeil P, Franklin CK, Fasquelle C, Géron S, Isacke CM, et al. Allelic heterogeneity of crooked tail syndrome: Result of balancing selection? Anim Genet 2012;43:604–7. doi:10.1111/j.1365-2052.2011.02311.x.

[23] Meydan H, Yildiz MA, Agerholm JS, Windsor P, Agerholm J, Kehrli M, et al. Screening for bovine leukocyte adhesion deficiency, deficiency of uridine monophosphate synthase, complex vertebral malformation, bovine citrullinaemia, and factor XI deficiency in Holstein cows reared in Turkey. Acta Vet Scand 2010;52:56. doi:10.1186/1751-0147-52-56.

[24] Patel RK, Singh KM, Soni KJ, Chauhan JB, Sambasiva Rao KRS. Lack of carriers of citrullinaemia and DUMPS in Indian Holstein cattle. J Appl Genet 2006;47:239–42. doi:10.1007/BF03194629.

[25] Nissen PH, Shukri NM, Agerholm JS, Fredholm M, Bendixen C. Genetic mapping of spinal dysmyelination in cross-bred American Brown Swiss cattle

to bovine Chromosome 11. Mamm Genome 2001;12:180–2. doi:10.1007/s003350010245.

[26] Krebs S, Medugorac I, Röther S, Strässer K, Förster M. A missense mutation in the 3-ketodihydrosphingosine reductase *FVT1* as candidate causal mutation for bovine spinal muscular atrophy. Proc Natl Acad Sci USA 2007;104:6746–51. doi:10.1073/pnas.0607721104.

[27] Medugorac I, Kemter J, Russ I, Pietrowski D, Nüske S, Reichenbach HD, et al. Mapping of the bovine spinal muscular atrophy locus to Chromosome 24. Mamm Genome 2003;14:383–91. doi:10.1007/s00335-002-3024-3.

[28] Hirano T, Kobayashi N, Matsuhashi T, Watanabe D, Watanabe T, Takasuga A, et al. Mapping and exome sequencing identifies a mutation in the *IARS* gene as the cause of hereditary perinatal weak calf syndrome. PLoS One 2013;8. doi:10.1371/journal.pone.0064036.

[29] Hirano T, Matsuhashi T, Takeda K, Hara H, Kobayashi N, Kita K, et al. *IARS* mutation causes prenatal death in Japanese Black cattle. Anim Sci J 2016;87:1178–81. doi:10.1111/asj.12639.

[30] Duchesne A, Gautier M, Chadi S, Grohs C, Floriot S, Gallard Y, et al. Identification of a doublet missense substitution in the bovine *LRP4* gene as a candidate causal mutation for syndactyly in Holstein cattle. Genomics 2006;88:610–21. doi:10.1016/j.ygeno.2006.05.007.

[31] Drögemüller C, Leeb T, Harlizius B, Tammen I, Distl O, Höltershinken M, et al. Congenital syndactyly in cattle: four novel mutations in the low-density lipoprotein receptor-related protein 4 gene (*LRP4*). BMC Genet 2007;8:5. doi:10.1186/1471-2156-8-5.

[32] Randhawa IA, Khatkar M, Thomson P, Raadsma H. Composite selection signals can localize the trait specific genomic regions in multi-breed populations of cattle and sheep. BMC Genet 2014;15:34. doi:10.1186/1471-2156-15-34.

[33] Pohlkamp T, Durakoglugil M, Lane-Donovan C, Xian X, Johnson EB, Hammer RE, et al. *LRP4* domains differentially regulate limb/brain

development and synaptic plasticity. PLoS One 2015;10:1–11. doi:10.1371/journal.pone.0116701.

[34] Ruíz-Flores A, Núñez-Domínguez R, Ramírez-Valverde R, Domínguez-Viveros J, Mendoza-Domínguez M, Martínez-Cuevas E. Niveles y efectos de la consanguinidad en variables de crecimiento y reproductivas en bovinos Tropicarne y Suizo Europeo. Agrociencia 2006;40(3):289-301.

Tables

Table 1. Frequency of homozygotes and heterozygotes in the studied population for genetic diseases previously reported in Braunvieh/Brown Swiss and other economically important genetic diseases in cattle.

Disease	Associated Gene Name	Gene Symbol	Chr	Fq Ht[a]	Fq Ho[b]
Angus Dwarfism	Protein kinase cGMP-dependent type II	*PRKG2*	6	N/P	
α-Mannosidosis	Mannosidase alpha class 2B member 1	*MAN2B1*	7	N/P	
β-Mannosidosis	Mannosidase beta	*MANBA*	16	N/P	
Bovine arachnomelia syndrome	Molybdenum cofactor synthesis 1	*MOCS1*	23	N/P	
Bovine citrullinaemia	Argininosuccinate synthase 1	*ASS1*	11	N/P	0.003
Bovine spinal dysmyelination	Spastin	*SPAST*	11	0.003	0.003
Bovine spinal muscular atrophy	AFG3 like matrix AAA peptidase subunit 2	*AFG3L2*	24	N/P	0.003
Braunvieh fertility haplotype 2	Tubulin delta 1	*TUBD1*	19	N/P	0.003
Chediak Higashi syndrome	Lysosomal trafficking regulator	*LYST*	28	N/P	
Congenital muscular dystonia	ATPase sarcoplasmic/endoplasmic reticulum Ca2+ transporting 1	*ATP2A1*	25	0.005	0.003
Congenital myasthenic syndrome	Cholinergic receptor nicotinic epsilon subunit	*CHRNE*	19	N/P	
Crooked Tail Syndrome	Mannose receptor C type 2	*MRC2*	19	0.014	N/P
Dilated cardiomyopathy	Outer mitochondrial membrane lipid metabolism regulator	*OPA3*	18	N/P	
DUMPS[c]	Uridine monophosphate synthase	*UMPS*	1	0.003	N/P

Table 1. Frequency of homozygotes and heterozygotes in the studied population for genetic diseases previously reported in Braunvieh/Brown Swiss and other economically important genetic diseases in cattle (Continue).

Disease	Gene Name	Symbol	Chr	Fq Ht[a]	Fq Ho[b]
Epidermolysis bullosa	Laminin subunit gamma 2	*LAMC2*	16	N/P	0.003
Factor XI	Coagulation factor XI	*F11*	27		N/P
Glycogen storage disease II	Glucosidase alpha, acid	*GAA*	19	N/P	0.003
Glycogen storage disease V	Glycogen phosphorylase, muscle associated	*PYGM*	29		N/P
Hypotrichosis	Hephaestin like 1	*HEPHL1*	29	0.003	N/P
Marfan syndrome	Fibrillin 1	*FBN1*	10	0.003	N/P
Maple syrup urine disease	Branched chain keto acid dehydrogenase E1, alpha polypeptide	*BCKDHA*	18	0.003	0.003
Mule foot disease[d]	LDL receptor related protein 4	*LRP4*	15	0.355	0.237
Protoporphyria	Ferrochelatase	*FECH*	24		N/P
Pseudomyotonia	ATPase sarcoplasmic/endoplasmic reticulum Ca2+ transporting 1	*ATP2A1*	25		N/P
BLAD[e]	Integrin subunit beta 2	*ITGB2*	1		N/P
Chondrodysplastic dwarfism	EvC ciliary complex subunit 2	*EVC2*	6		N/P
Weak calf syndrome[f]	Isoleucyl-tRNA synthetase	*IARS*	8	0.003	N/P
Weaver syndrome[g]	Patatin like phospholipase domain containing 8	*PNPLA8*	4	0.008	0.007

[a]Fq Ht: frequency of heterozygotes. [b]Fq Ho: frequency of homozygotes. N/P: animals with disease associated markers not present in the population. [c]Deficiency of uridine monophosphate synthase.[d]Syndactyly syndrome; [e]Bovine leukocyte adhesion deficiency; [f]Hereditary perinatal weak calf syndrome; [g]Bovine progressive degenerative myeloencephalopathy.

Table 2. Allele and genotype frequency of markers previously associated with genetic diseases present in the studied population.

Disease	Gene Symbol	Marker Name	Marker position	Chr	SNP*	Allele*		Genotype		
						1	2	11	12	22
BMS[a]	*FBN1*	FBN1_1	62054844	10	G/A	0.998	0.002	0.997	0.003	n/p
BSD[b]	*SPAST*	SDM	n/e	11	G/A	0.997	0.003	0.997	n/p	0.003
		SDM_2				0.998	0.002	0.997	0.003	n/p
		SDM_3				0.997	0.003	0.997	n/p	0.003
BH2[c]	*TUBD1*	BH2	n/e	24	T/C	0.997	0.003	0.997	n/p	0.003
Citrullinaemia	*ASS1*	ASS1 Citrullinemia_3	100802781 n/e	11	C/T	0.997	0.003	0.997	n/p	0.003
CMD[d]	*ATP2A1*	CMD1	n/e	25	T/C	0.997	0.003	0.997	n/p	0.003
		CMD1_3				0.997	0.003	0.993	0.007	n/p
		CMD2_2			A/G	0.998	0.002	0.997	0.003	n/p
Crooked Tail	*MRC2*	MRC2_1	47734925	19	I/D	0.984	0.006	0.969	0.031	n/p
		CTS-BB	47740444			0.997	0.003	0.993	0.007	n/p
		CTS-BB_2	0			0.995	0.005	0.990	0.010	n/p
		CTS-BB_3				0.997	0.003	0.993	0.007	n/p
DUMPS[e]	*UMPS*	DUMPS	n/e	1	C/T	0.998	0.002	0.997	0.003	n/p
		DUMPS_2								
		DUMPS_3								

Table 2. Allele and genotype frequency of markers previously associated with genetic diseases present in the studied population (Continue I).

Disease	Gene Symbol	Marker Name	Marker position	Chr	SNP*	Allele*		Genotype		
						1	2	11	12	22
EB[f]	*LAMC2*	EB_2	4164**	5	G/A	0.998	0.002	0.997	n/p	0.003
Hypotrichosis	*HEPHL1*	HEPHL1	695072	29	A/T	0.998	0.002	0.997	0.003	n/p
Mulefoot[g]	*LRP4*	LRP4_3	77675440	15	C/A	0.705	0.295	0.507	0.396	0.097
		Mulefoot-241	Exon 3		G/A	0.483	0.517	0.253	0.461	0.286
		Mulefoot-241_2	Exon 3			0.487	0.513	0.256	0.461	0.283
		Mulefoot-241_3	Exon 3			0.490	0.510	0.263	0.454	0.283
		Mulefoot-2719	Exon 20			0.998	0.002	0.997	0.003	n/p
MSUD[h]	*BCKDHA*	MSUD	n/e	18	C/T	0.998	0.002	0.997	0.003	n/p
		MSUD_2				0.997	0.002	0.997	n/p	0.003
		MSUD_3				0.998	0.003	0.997	0.003	n/p
SMA[i]	*AFG3L2*	SMA	Distal part of Chr24	24	G/T	0.997	0.003	0.997	n/p	0.003
Pompes[j]	*GAA*	Pompes_1783_BR	Exon 13	19	C/T	0.997	0.003	0.997	n/p	0.003
		Pompes_1783_BR_2	Exon 13							
		Pompes_1783_BR_3	Exon 13							
WCS[k]	*IARS*	IARS	85341291	8	G/C	0.998	0.002	0.997	0.003	n/p

Table 2. Allele and genotype frequency of markers previously associated with genetic diseases present in the studied population (Continue II).

Disease	Gene Symbol	Marker Name	Marker position	Chr	SNP*	Allele*		Genotype		
						1	2	11	12	22
WS[I]	*PNPLA8*	WVR_49656945	49656945	4	C/T	0.996	0.004	0.993	0.007	n/p
		WVR_49657798	49657798		G/A	0.989	0.011	0.986	0.007	0.007
		WVR_49664852	49664852		C/T	0.996	0.004	0.993	0.007	n/p
		WVR_49667361	49667361		C/T	0.996	0.004	0.993	0.007	n/p
		WVR_49673503	49673503		C/T	0.993	0.007	0.987	0.013	n/p
		WVR_49681169	49681169		C/T	0.996	0.004	0.993	0.007	n/p
		WVR_49682552	49682552		A/G	0.996	0.004	0.993	0.007	n/p
		WVR_49686038	49686038		A/G	0.996	0.004	0.993	0.007	n/p
		WVR_49687224	49687224		T/G	0.996	0.004	0.993	0.007	n/p
		WVR_49691015	49691015		G/T	0.989	0.011	0.986	0.007	0.007
		WVR_49692015	49692015		T/C	0.996	0.004	0.993	0.007	n/p
		WVR_49692485	49692485		G/T	0.989	0.011	0.986	0.007	0.007
		WVR_49692825	49692825		C/T	0.990	0.010	0.980	0.020	n/p
		WVR_49693121	49693121		C/T	0.996	0.004	0.993	0.007	n/p
		WVR_49693140	49693140		G/T	0.993	0.007	0.987	0.013	n/p
		WVR_49693164	49693164		A/T	0.996	0.004	0.993	0.007	n/p

Table 2. Allele and genotype frequency of markers previously associated with genetic diseases present in the studied population (Continue III).

Disease	Gene Symbol	Marker Name	Marker position	Chr	SNP	Allele		Genotype		
						1	2	11	12	22
		WVR_49693265	49693265		A/G	0.996	0.004	0.993	0.007	n/p
		WVR_49693601	49693601		T/C	0.996	0.004	0.993	0.007	n/p
		WVR_49693913	49693913		A/C	0.996	0.004	0.993	0.007	n/p
		WVR_49694977	49694977		A/G	0.996	0.004	0.993	0.007	n/p
		WVR_49695504	49695504		A/G	0.993	0.007	0.987	0.013	n/p
		WVR_49695952	49695952		C/T	0.989	0.011	0.986	0.007	0.007
		WVR_49698002	49698002		C/T	0.993	0.007	0.987	0.013	n/p
WS	*PNPLA8*	WVR_49698436	49698436	4	C/T	0.996	0.004	0.993	0.007	n/p
		WVR_49700154	49700154		A/G	0.996	0.004	0.993	0.007	n/p
		WVR_49701106	49701106		T/C	0.993	0.007	0.987	0.013	n/p
		WVR_49702287	49702287		T/C	0.996	0.004	0.993	0.007	n/p
		WVR_49702494	49702494		G/A	0.996	0.004	0.993	0.007	n/p
		WVR_49708283	49708283		T/C	0.996	0.004	0.993	0.007	n/p
		WVR_49715678	49715678		C/T	0.996	0.004	0.993	0.007	n/p
		WVR_49718641	49718641		C/T	0.996	0.004	0.993	0.007	n/p

aBMS: bovine Marfan syndrome; bBSD: bovine spinal dysmyelination; cBH2: Braunvieh fertility haplotype 2; dCMD: congenital muscular dystonia; eDUMPS; deficiency of uridine monophosphate synthase. fEB: epidermolysis bullosa; gMulefoot: syndactyly syndrome; hMSUD: maple syrup urine disease. iSMA: bovine spinal muscular atrophy; jPompes: glycogen storage disease II; kWCS: hereditary perinatal weak calf syndrome. lWS: bovine progressive degenerative myeloencephalopathy (Weaver syndrome). n/e: not specified in the microarray. n/p: genotype not present in the population. *The first letter in the SNP column represents the allele 1 and the second letter of the same column represents the allele 2; Allele 2 is the causing variant of the disease **Position in the GenBank accession number AY740402. n/p: genotype not present in the population.

6. GENOME-WIDE ASSOCIATION ANALYSIS FOR GROWTH TRAITS IN MEXICAN BRAUNVIEH CATTLE

J. L. Zepeda-Batista[*], G. M. Parra-Bracamonte†, R. Núñez-Domínguez[*], R. Ramírez-Valverde[*], A. Ruíz-Flores[*], F. J. Jahuey-Martínez¶,

[*]Universidad Autónoma Chapingo. Departamento de Zootecnia. Posgrado en Producción Animal. Km. 38.5 Carretera México-Texcoco. Chapingo, Estado de México. C.P. 56230

†Laboratorio de Biotecnología Animal. Centro de Biotecnología Genómica. Instituto Politécnico Nacional. Boulevard del Maestro S/N esq. Elías Piña, Col. Narciso Mendoza, Ciudad Reynosa, Tamaulipas, México, C.P. 88710.

¶Facultad de Zootecnia y Ecología, Universidad Autónoma de Chihuahua. Periférico Francisco R. Almada. Km 1. Chihuahua, Chihuahua, México. C.P. 33820.

Submitted to the *Journal of Animal Science*

Genome-wide association analysis for growth traits in Mexican Braunvieh cattle

J. L. Zepeda-Batista*, G. M. Parra-Bracamonte†, ¶, R. Núñez-Domínguez*, R. Ramírez-Valverde*, A. Ruíz-Flores*, F. J. Jahuey-Martínez

*Universidad Autónoma Chapingo. Departamento de Zootecnia. Posgrado en Producción Animal. Km. 38.5 Carretera México-Texcoco. Chapingo, Estado de México. C.P. 56230

†Centro de Biotecnología Genómica, Instituto Politécnico Nacional. Boulevard del Maestro S/N esq. Elías Piña, Col. Narciso Mendoza. Ciudad Reynosa, Tamaulipas, México, C.P. 88710.

¶Facultad de Zootecnia y Ecología, Universidad Autónoma de Chihuahua. Periférico Francisco R. Almada. Km 1. Chihuahua, Chihuahua, México. C.P. 33820.

Correspondence author

Gaspar Manuel Parra Bracamonte; E-mail: gparra@ipn.mx

ABSTRACT: The study aimed to perform a genome-wide association study (GWAS) for growth weight traits in Braunvieh cattle to identify SNP markers and genes associated with these traits. The study comprised 300 animals genotyped, using 30,125 SNP markers from the GeneSeek® Genomic Profiler Bovine LDv.4 panel. The examined phenotypic data included birth, weaning, and yearling weights. After quality control, 22,734 SNP and 276 animals were retained in the final analysis. The association analysis was performed using the principal components method, via the egscore function of the GenABEL version 1.8-0 package in the R environment. The marker rs133262280 located in BTA 22 was associated with birth weight, while there were two SNPs associated with weaning weight, rs43668789 (BTA 11) and rs136155567 (BTA 27). New QTL were detected in association with the growth traits and identified 4 positional and functional candidate genes potentially involved with a variation of the analyzed traits. The most important genes in these genomic regions were *MCM2* (minichromosome maintenance complex component 2), *TPRA1* (transmembrane protein adipocyte associated 1), GALM (*galactose mutarotase*), and *NRG1* (neuregulin 1), due to their relationships with embryonic cleavage, bone and tissue growth, cell adhesion, and organ development. This study is the first to describe a GWAS conducted in Braunvieh cattle in Mexico and represents a benchmark for future research with this breed. Further analyses of these regions could help to identify useful markers for marker-assisted selection and will contribute to the knowledge of the genetic basis of growth in cattle and be a basis for genomic prediction assessment in Mexican Braunvieh cattle.

Key words: *Bos taurus*, quantitative trait loci, single nucleotide polymorphism, population structure

INTRODUCTION

Since the beginning of the beef production industry, beef cattle have been genetically improved by the breeders. The most common traits improved in the meat production industry have been those related with growth, and to a lesser degree feeding, reproduction, and meat quality (Chin-Colli *et al*., 2016; Jahuey-Martínez *et al*., 2016). One worldwide cattle breed used in the beef industry is the Braunvieh, in specialized beef, dairy or dual-purpose production systems (Phillips *et al*., 2009; Orantes-Zebadúa *et al*., 2014). Due to its initial dual-purpose origin, most of the available information about Braunvieh deals with dairy traits, although lately, Braunvieh cattle has been studied for beef production traits (Phillips *et al.*, 2009; Chin-Colli *et al*., 2016), since growth traits are good indicators of animal productivity, viability and efficiency of farms engaged in meat production. Birth, weaning, and yearling weights, have been studied in Mexican Braunvieh with little genetic progress (Chin-Colli *et al*., 2016).

Recently, many QTL affecting production traits in beef cattle were located (Rolf *et al.*, 2011; Purfield *et* al., 2015; Jahuey *et al*., 2016) but most of the association studies focused on specialized beef breeds and only a few types of research have been implemented in the minor, but still global, breeds such as Braunvieh (Guo *et al*., 2012; Maxa *et al*., 2012). The use of thousands of SNP markers in genome-wide association studies (GWAS) has allowed the discovery and confirmation of many QTL for growth traits in beef and crossbred cattle (Lu *et al.*, 2013; Jahuey-Martínez *et al.*, 2016; Martínez *et al*., 2016), which in turn have served as the basis for the search of the nucleotides responsible for the phenotypic variation (Takasuga, 2016).

In Mexico, Braunvieh is one of the cattle breeds most used in the beef production industry, either as purebred or in crosses with *Bos indicus* cattle (AMCGSR, 2017, Orantes-Zebadúa *et al*., 2014). Despite the long time that Braunvieh cattle has been used in Mexico, there is scarce information on the breed productive performance, and the available information is mostly related with growth traits coming from genetic evaluations or isolated studies (Silva *et al*., 2002; Chin-Colli *et al*., 2016; AMCGSR, 2017). Therefore, the selection, management, and genetic improvement programs of the Braunvieh cattle could potentially benefit from the use of high-throughput genotyping technologies.

This study aimed to perform a GWAS, using genome-wide SNP markers in the Mexican registered Braunvieh cattle population, to identify QTLs for growth traits and to define genes as potential candidates for further studies.

MATERIALS AND METHODS

Approval from the ethical committee for animal care and use was not necessary because the samples used in this study consisted of hair follicles.

Population and phenotypic data

Hair follicles samples were collected from 236 females and 64 males registered in the database of the Mexican Braunvieh Cattle Association. The cattle were born between 2000 and 2015. This population came from herds located in the east, west, and central highlands of Mexico. Herds from west and east were raised under extensive production systems, whilst central highlands herds were under intensive regimen. Genetic background of the sampled population included Austrian, Swiss, Canadian, American and Mexican animals. Phenotypic data were provided by the breeding association and included records of birth

(BWT, kg), weaning (WW, kg), and yearling weights (YW, kg). Weaning and yearling weights were adjusted to 205 days and 365 days, respectively, according to the guidelines of the Beef Improvement Federation (2016) to use in the GWAS analysis. Table 1 shows the descriptive statistics for each trait.

Table 1. Descriptive statistics for growth traits of Mexican Braunvieh cattle

Trait[1]	*N*	*n*(QC)[2]	Mean	SD	Minimum	Maximum
BWT	300	266	38.01	4.07	22	50
WW	300	263	212.40	27.43	128	308
YW	300	244	313.17	45.47	176	440

[1]BWT: birth weight; WW: weaning weight; YW = yearling weight.
[2]*n*(QC) = *n* after quality control.

Genotyping and quality control

The animals were genotyped using 30,125 SNP markers from the GeneSeek® Genomic Profiler Bovine LDv.4 panel (Neogen Corp. Lincoln, NE, USA). Before association analysis, the quality of the genotypic data was verified using the SNPQC program (Gondro *et al*., 2014). Animals were eliminated if they exhibited call rates of less than 80% ($n = 0$) or levels of heterozygosity (HE) above 3 SD ($n = 1$), considering that the mean and SD of the observed HE was 0.32 and 0.019, respectively.

Genotypes were considered successful if they presented a GenCall value greater than 0.50, and all SNPs with lower values were discarded ($n = 1623$). Those SNPs that were monomorphic ($n = 3604$), presented call rates lower than 90% ($n = 1290$), minor allele frequencies < 0.01 ($n = 1325$), or deviated from Hardy-Weinberg equilibrium according to Fisher's exact test and exhibited *P*-values > 1×10^{-15} ($n = 0$) were also eliminated. In addition, SNPs with unknown coordinates in the assembly of the bovine genome UMD

v3.1 (Zimin *et al*., 2009) (n = 1484) and SNPs that were not located on autosomal chromosomes (n = 1820) were discarded.

Finally, a Pearson correlation was computed between each pair of samples according to their genotype information, obtaining an average of $r = 0.817$ and minimum and maximum values of 0.663 and 0.900, respectively. A maximum value of $r \geq 0.98$ for detecting potentially duplicate samples (n = 0) was also considered. A total of 22,734 SNPs and 276 samples passed the quality control procedures and were retained for further analysis. Quality control and subsequent analyses were performed in the R environment.

Population structure and association analysis

Population structure was analyzed calculating first a genomic relationship matrix, using information on genotypes as suggested by VanRaden (2008), besides performing a singular value decomposition and a principal components (PC) analysis.

The analysis of PCs indicated that 28.6% of the variance in the data was explained by the first two components. Therefore, it was decided to perform a genome-wide association analysis using the PC method proposed by Price *et al*. (2006). For this analysis, the egscore function from the GenABEL package of R (Aulchenko *et al*., 2007) was used. To account for population stratification, this function uses the genomic kinship matrix to derive axes of genetic variation and then, both the phenotypes and genotypes are adjusted onto these axes.

A linear model for each trait was fitted including the first two PC as covariates. For the analysis of BWT, the model also included the contemporary group (CG) and the linear and quadratic effects of age of the dam when birth and weaning weight were measured.

The CG included herd, sex, year and season of birth. The statistical models used to analyze the other traits only included the CG as well as the PCs as covariates, the age of dam was excluded from the final model because it was not a significant factor in the preliminary analysis. Finally, the association between corrected genotypes and phenotypes was assessed via correlation, and *P*-values were obtained by calculating the square of the correlation multiplied by (N-K-1), where N was the number of genotyped individuals, and K was the number of PCs.

Minimum allele frequencies, allele substitution effect (β) and percentage of phenotypic variance explained by the SNP were estimated. The proportion of phenotypic variance explained by the SNPs was estimated by dividing the X^2 value for a df by the number of individuals used for the analysis of each SNP marker, followed by multiplication by 100. SNPs with P-values $< 5 \times 10^{-5}$ were considered significantly associated with the studied traits. All described analyses and estimations were performed using GenABEL package (Aulchenko *et al*., 2007).

Analysis of genomic regions with significant SNPs

The closest genes to significant markers and those located within a 250-kb window on both sides of the SNP location were identified. The list of genes was obtained using the snp2gene.LD function from the Postgwas package (Hiersche *et al*., 2014). The distance between SNPs and genes was calculated as the difference between the marker position and the beginning or end of the gene according to the coordinates from the assembly of the bovine genome UMD v3.1. Gene functions were investigated in the UniProt database (UniProt Consortium, 2017).

Annotations from humans or mice were used when there was no information on the genes in cattle. Genes were considered functional and positional candidates if they were biologically related to the trait under study, supported by experimental evidence in literature. Finally, it was determined if significant SNPs mapped against QTLs were previously associated with growth-related traits such as BWT, carcass and reproduction traits, and deposited or not in the cattle AnimalQTLdb (Hu *et al.*, 2013). For this purpose, we used SNP positions according to the Btau4.6 genome sequence because many of the previously reported QTLs had no well-defined positions in the assembly of the bovine genome UMD v3.1.

RESULTS AND DISCUSSION

A total of 30,125 SNP markers from the GeneSeek® Genomic Profiler Bovine LD v4 panel (Neogen Corp. Lincoln, NE, USA) were evaluated for associations with growth traits of Braunvieh cattle. On average, 1,004 SNP markers were evaluated in each BTA. *Bos taurus* chromosomes 1 and 27 exhibited the highest (1602) and lowest (512) numbers of SNP, respectively. The average distance between adjacent SNPs was 87,641 bp, the minimum distance (0 bp) between adjacent SNP was found on BTA 1, 6, 7, 12, 17, 18, 22, 25, 26, 28, and 29 while the maximum distance (1,962,000 bp) was found on BTA 6.

The results showed the presence of only two genetic populations well differentiated (Figure 1). These results were expected by the fact that the tested herds presented different selection objectives and ancestors of the imported germplasm (*i. e.*, semen, sires). Stratification of results could attend to an extensive use of sires or semen that breeders usually choose in their genetic improvement programs.

Several authors (Harris *et al*., 2010; Erbe *et al*., 2012; Plieschke *et al*., 2015) have performed the study of subdivisions aimed to detect the effect of those subdivisions on the genomic estimated breeding value (GEBV) and the estimation of QTLs using genome wide association studies (GWAS). On this regard, Smitz *et al.* (2014) concluded that stratification in the studied populations needs to be considered in genetic improvement programs to conserve the "genetic health" of those populations.

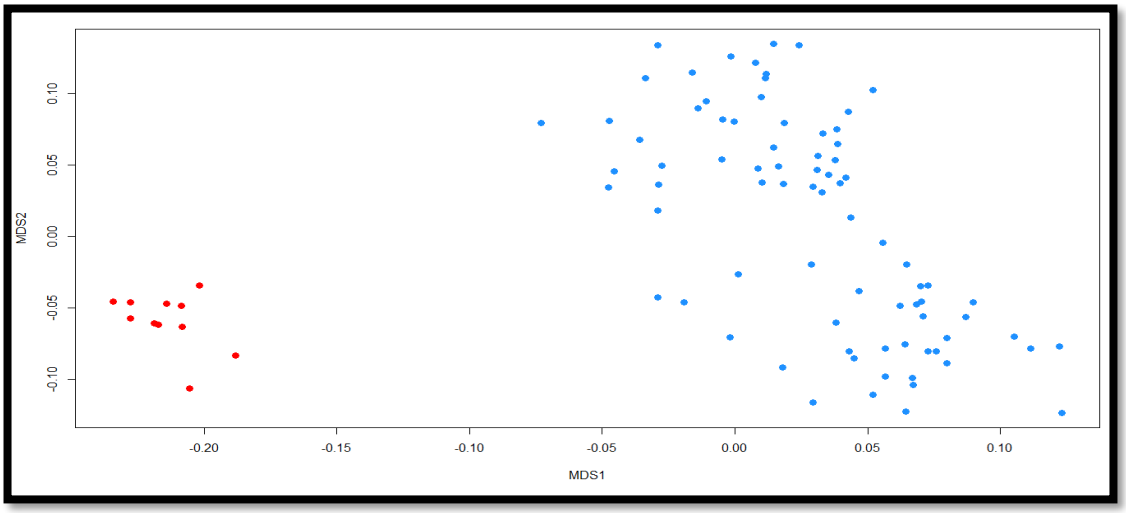

Figure 1. Presence of two subpopulations in the sample of the Braunvieh population analyzed. Circles of different color indicate different population. MDS1: multi-dimension scalling 1; MDS2: multi-dimension scalling 2.

Ben Jemaa *et al*. (2015) indicate that some QTLs found in GWAS could not be present in all the studied animals due to the stratification of the population. Aiming to obtain representativeness, the effect of population structure was considered in this study and the results could be observed in the quantile-quantile plot for each GWAS (Fig. 2)

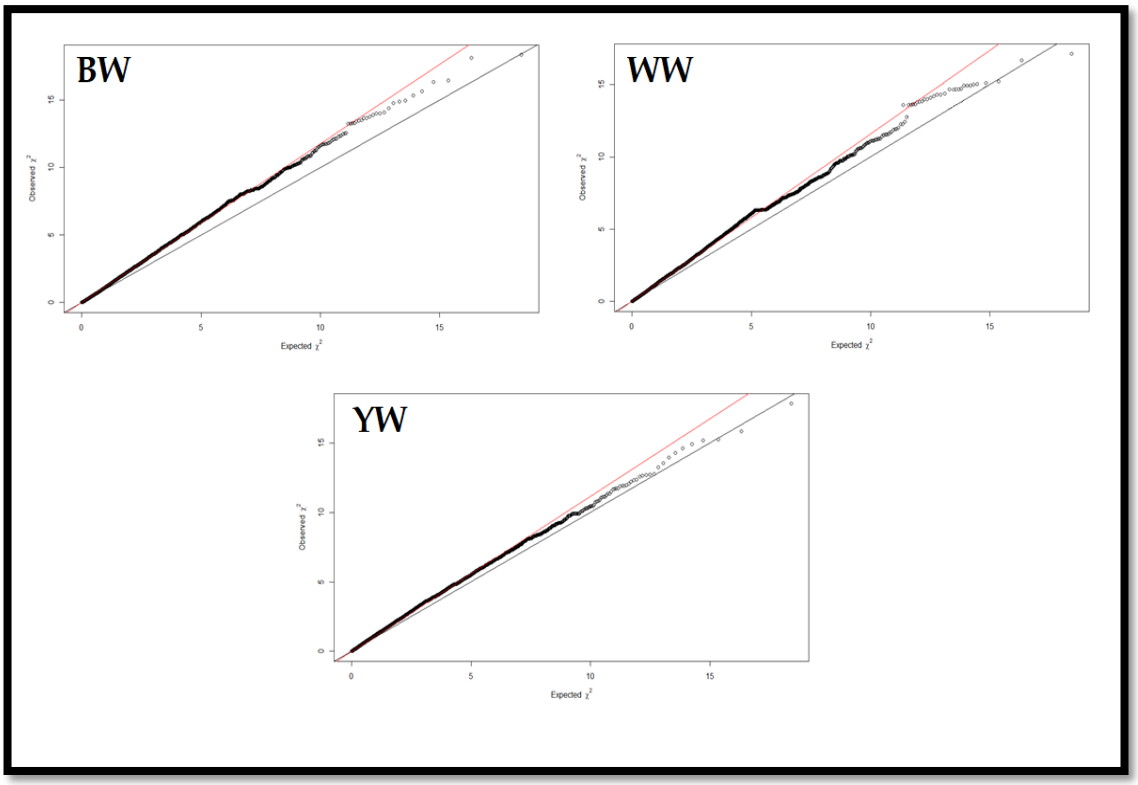

Figure 2. Q-Q plot of growth traits Quantile-quantile (QQ) plots for the genome wide association study of growth traits in Braunvieh cattle. The straight line in the QQ plots indicates the distribution of SNP markers under the null hypothesis, and the skew at the edge indicates that these markers are more strongly associated with the traits than would be expected by chance. BWT = birth weight; WW = weaning weight; YW = yearling weight.

According to the significance threshold considered ($P < 5 \times 10^{-5}$), 3 SNP were associated with the growth traits (Table 2). These markers were distributed on BTA 11, 22, and 27. Figure 3 shows the Manhattan plots in which the −log10 transformations of the P-values were plotted for each of the GWAS. Genes and QTL previously associated with growth-related traits are shown in Table 3. Tables 4 to 6 show complete descriptions including the identifier number and exact location of each gene as well as any previously reported QTL located in the genomic regions identified in this study.

Table 2. Parameters and statistics of SNP associated with growth traits of Mexican Braunvieh cattle[1]

Trait	SNP ID[2]	BTA	UMD3.1,[3] bp	Btau4.6,[4] bp	Allele	MAF[5]	β,[5] kg	SE	Percentage Var[5]	*P*-value
BWT	rs133262280	22	60,759,211	127,745,473	C/T	0.18	0.320	0.02	0.1	0.0000274
WW	rs43668789	11	21,312,462	22,502,811	C/T	0.17	-9.590	0.25	2.98	0.0000528
	rs136155567	27	27,056,807	29,944,194	A/G	0.20	1.110	0.72	1.1	0.0000127

[1]BWT = birth weight; WW = weaning weight.
[2]ID = identification.
[3]UMD version 3.1 (Zimin *et al*., 2009).
[4]Elsik *et al*. (2015).
[5]MAF = minimum allele frequency; β = allele substitution effect; Var = phenotypic variance explained by the SNP.

The present study identified two regions (WW_rs43668789_11_21.3 and WW_rs136155567_27_27.0) previously reported by McClure *et al*. (2010) as associated with weaning weight and calving ease in Angus cattle. Besides, Boichard *et al.* (2003) and Buitenhuis *et al*. (2007) reported associations between the identified regions in this study and conformation traits, explaining between 5.9 to 8.9 % of the structural soundness in 10 European dairy cattle breeds. On the other hand, Sherman *et al.* (2009) and Rolf *et al*. (2012) reported associations with allele substitution effects between -0.319 to 2.199 kg for feeding traits like average daily gain and residual feed intake in Angus, Charolais, and Canadian beef hybrid cattle.

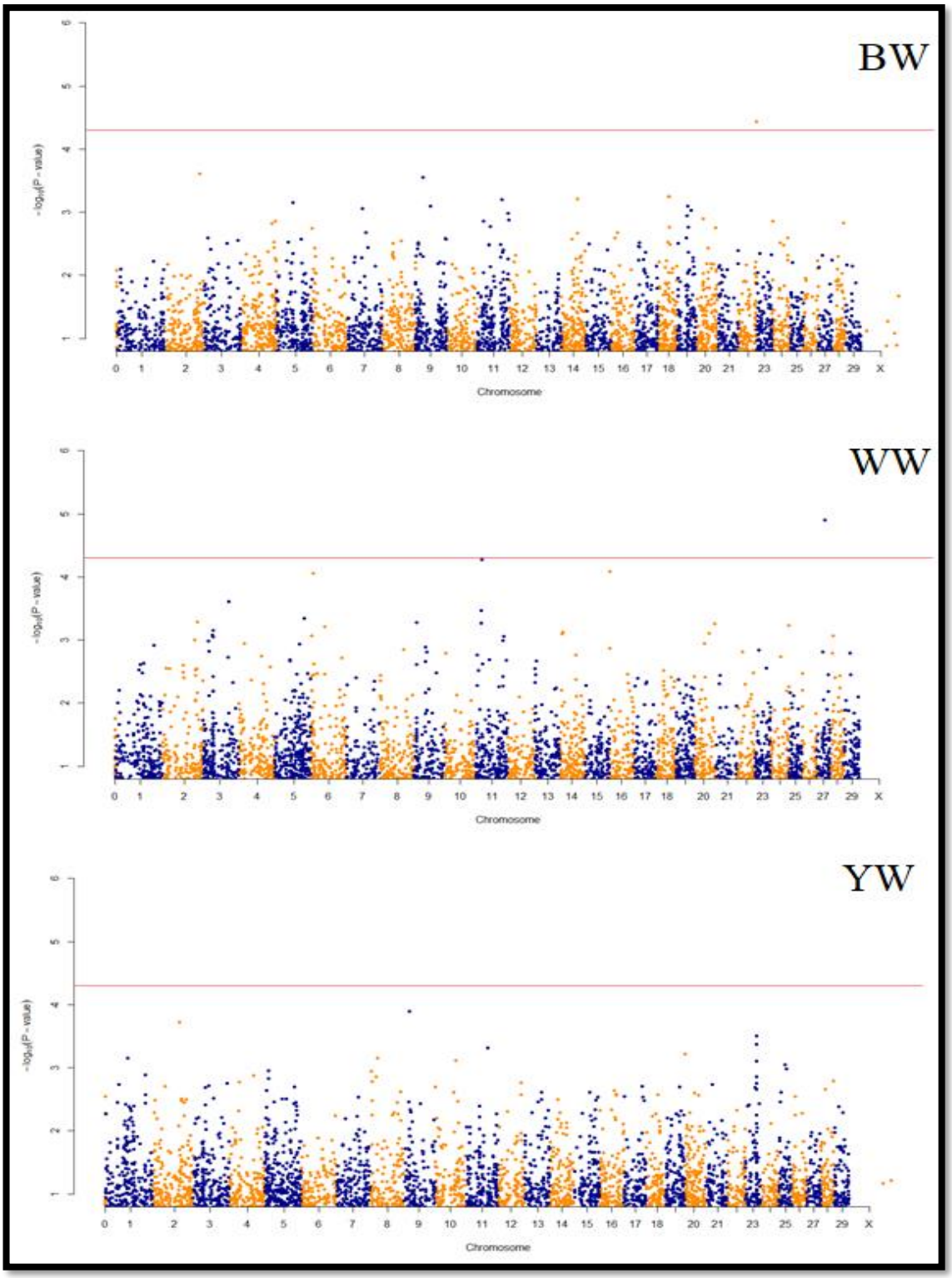

Figure 3. Manhattan plots of the *P*-values for the genome-wide association study of growth traits in Braunvieh cattle. The horizontal line indicates the significance threshold for significant associations ($P < 5 \times 10^{-5}$). BWT = birth weight; WW = weaning weight; YW = yearling weight.

Table 3. Genes and previously reported QTL[1] located near significant associated SNP

Trait_SNP ID[2]_BTA_Mb	Genes in ±250 kb	QTL	Reference
BWT_rs133262280_22_60.7	PODXL2, MCM2, TPRA1, LOC10105309, PLXNA1, CHCHD6	-	-
WW_rs43668789_11_21.3	ARHGEF33, MORN2, DHX57, GEMIN6, SRSF7, GALM, MIR2284Z-2, SOS1, CDKL4, MAP4K3	SOUND	Buitenhuis *et al*., 2007
		RFI	Sherman *et al*., 2009
		RANGLE	Boichard *et al*., 2003
		WWMM	McClure *et al*., 2010
WW_rs136155567_27_27.0	LOC104976093, NRG1	BQ	Buitenhuis *et al*., 2007
		SOUND	Buitenhuis *et al*., 2007
		ADFI	Rolf *et al*., 2012
		ADG	Rolf *et al*., 2012
		RFI	Rolf *et al*., 2012
		CALEASE	McClure *et al*., 2010

[1]ADG= average daily gain; ADFI = average daily feed intake; BQ = bone quality; BWT = birth weight; CALEASE = calving ease; RFI = residual feed intake; RANG = rump angle; SOUND = structural soundness; WW = weaning weight; WWMM = weaning weight-maternal milk.

[2]ID = identification.

Birth weight

Birth weight in Braunvieh cattle represents an important trait to consider in the genetic improvement programs, due to its association with calving difficulty in young heifers, especially when Braunvieh is used as a sire with smaller breeds (Hagger and Hofer, 1990). The present study identified the rs133262280 as the only marker associated with BWT and was located at 60.7 Mb of BTA 22. This SNP showed an allelic substitution effect of

0.320 kg, explaining 0.1% of the phenotypic variance of BWT. Genes located closer to this SNP included CHCHD6 (coiled-coil-helix-coiled-coil-helix domain containing 6), LOC10105309 (uncharacterized LOC101905309), MCM2 (minichromosome maintenance complex component 2), PLXNA1 (plexin A1), PODXL2 (podocalyxin like 2), and TPRA1 (transmembrane protein adipocyte associated 1) (Figure 4-Supplements). The most important genes identified in this region were *MCM2* and *TPRA1*, the first one is located at 177.6 kb, whilst *TPRA1* is just at 160.1 kb, both genes are upstream of the rs133262280 SNP. *MCM2* acts as a component of the MCM2-7 complex (MCM complex) which is the putative replicative helicase essential for 'once per cell cycle' DNA replication initiation and elongation in eukaryotic cells (Todorov *et al*., 1994). Additionally, this gene plays a role in cell division and apoptosis (Gao *et al*., 2015). Gao *et al*. (2015) reported MCM2 protein expression in the cochlea of rats and guinea pigs slight increase the apoptosis rate of the cells without any changes in proliferation or cell cycle.

Table 4. Genes close to the SNP rs133262280_22 associated with the birth weight of Mexican Braunvieh cattle

Gene in ±250 kb[1]	Gene ID[2]	Distance,[3] kb	Description
PODXL2	532521	U 202.2	Podocalyxin like 2
MCM2	510120	U 177.6	Minichromosome maintenance complex component 2
TPRA1	617772	U 160.1	Transmembrane protein adipocyte associated 1
LOC10105309	109905309	U 57.8	Uncharacterized LOC101905309
PLXNA1	531240	D 192.2	Plexin A1
CHCHD6	615934	D 200.9	Coiled-coil-helix-coiled-coil-helix domain containing 6

[1]rs136155567: gene in ±600 kb
[2]ID = identification.
[3]D = downstream; U = upstream.

The other associated gene with biological importance was TPRA1 gene, which belongs to the G protein-coupled receptor (GPCR) family. Functions related to this gene include the regulation of early embryonic cleavage and enhancing the hedgehog signaling pathway (Aki *et al*., 2008; Singh *et al*., 2015).

Table 5. Genes close to the SNPs associated to weaning weight of Mexican Braunvieh cattle.

SNP_BTA	Gene in ±250 kb[1]	Gene ID[2]	Distance,[3] kb	Description
rs43668789_11	GALM	616676	U 217.4	Galactose mutarotase
	SRSF7	507066	U 201.6	Serine and arginine rich splicing factor 7
	GEMIN6	525263	U 160.6	Gem nuclear organelle associated protein 6
	LOC107132913	107132913	U 156.0	Uncharacterized LOC107132913
	DHX57	540993	U 86.1	Dexh-box helicase 57
	MORN2	616607	U 77.8	MORN repeat containing 2
	ARHGEF33	100335703	Cover	Rho guanine nucleotide exchange factor 33
	SOS1	537682	D 17.0	SOS Ras/Rac guanine nucleotide exchange factor 1
	MIR2284Z-2	102465308	D 62.5	Microrna 2284z-2
	LOC104973309	104973309	D 121.0	Ubiquitin-40S ribosomal protein S27a pseudogene
	CDKL4	517478	D 207.4	Cyclin dependent kinase like 4
	LOC782845	782845	D 241.7	60S ribosomal protein L23a pseudogen
rs136155567_27	LOC104976093	104976093	D 470.9	Uncharacterized LOC104976093
	NRG1	281361	D 567.1	Neuregulin 1

[1]rs136155567: gene in ±600 kb
[2]ID = identification.
[3]D = downstream; U = upstream.

Several studies have highlighted its importance in pre- and perinatal tissue development in mice. Aki *et al.* (2015) determined that TPRA1 gene influenced the Hedgehog signaling pathway which plays an essential role in vertebrate embryonic tissue patterning of many developing organs, showing differences of around 50% in the signaling levels comparing homozygotes and heterozygotes animals. This evidence suggests that MCM2 and TPRA1 could participate in the early stages of development in cattle and could, therefore, influence BWT. There were any quantitative trait loci previously located in this region which indicates that could be a specific QTL of the studied population.

Weaning weight

Two SNP markers were associated with WW. One of these markers was rs43668789, located at 21.3 Mb of BTA 11 and showed an allelic substitution effect of -9.590 kg, explaining 2.98% of the phenotypic variance of WW. Genes located closer or covering this SNP included ARHGEF33 (Rho guanine nucleotide exchange factor 33), CDKL4 (cyclin dependent kinase like 4), DHX57 (DExH-box helicase 57), GALM (galactose mutarotase), GEMIN6 (gem nuclear organelle associated protein 6), LOC104973309 (ubiquitin-40S ribosomal protein S27a pseudogene), LOC107132913 (uncharacterized LOC107132913), LOC782845 (60S ribosomal protein L23a pseudogen), MAP4K3 (mitogen-activated protein kinase kinase kinase kinase 3), MIR2284Z-2 (microRNA 2284z-2), MORN2 (MORN repeat containing 2), SOS1 (SOS Ras/Rac guanine nucleotide exchange factor 1), and SRSF7 (serine and arginine rich splicing factor 7) (Figure 5-Supplements).

Table 6. Previously reported QTL[1] found near the SNP associated with growth traits of Mexican Braunvieh cattle

Trait_SNP ID2_BTA_Mb	QTL	QTL ID	QTL in Btau4.6,[3] bp	QTL reference
BWT_rs133262280_22_60.7	-	-	-	-
WW_rs43668789_11_21.3	SOUND	3591	18,215,471-23,417,727	Buitenhuis *et al*., 2007
	RFI	5281	8,076,786-33,430,175	Sherman *et al*., 2009
	RANGLE	3447	16,291,959-80,096,141	Boichard *et al*., 2003
	WWMM	10894	16,291,959-80,096,141	McClure *et al*., 2010
WW_rs136155567_27_27.0	BQ	3598	24,473,016-31,018,770	Buitenhuis *et al*., 2007
	SOUND	3594	24,473,016-31,018,770	Buitenhuis *et al*., 2007
	ADFI	21028	27,034,490-29,073,970	Rolf *et al*., 2012
	ADG	20979	27,034,490-29,073,970	Rolf *et al*., 2012
	RFI	21095	27,034,490-29,073,970	Rolf *et al*., 2012
	CALEASE	11259	21,801,052-31,012,980	McClure *et al*., 2010

[1]ADG= average daily gain; ADFI = average daily feed intake; BQ = bone quality; CALEASE = calving ease; RFI = residual feed intake; RANG = rump angle; SOUND = structural soundness; WWMM = weaning weight-maternal milk.
[2]ID = identification.

The most important gene identified in this region was *GALM*. This gene is located 217.4 kb upstream of the rs43668789 and belongs to the group of proteins that converts the α-aldose to β-anomer. *GALM* is involved in the pathway hexose metabolism, which is part of the carbohydrate metabolism (Thoden *et al*., 2004). McClure *et al.* (2010) reported a positive association of *GALM* with the weaning weight in Angus cattle. Besides, Shin *et al*. (2014) mentioned that the association between *GALM* and the weaning weight in Holstein and Hanwoo cattle lies in the quantity and the quality of the milk that the calves consume. Quantitative trait loci located in this region have been previously associated with weaning weight in Angus (McClure *et al*., 2010), conformation in dairy cattle breed (Boichard *et al.*, 2003; Buitenhuis *et al*., 2007), and residual feed intake in Canadian beef synthetic cattle (Sherman *et al*., 2009). The second marker associated with WW was rs136155567, located at 27.0 Mb of BTA 27 and its allele substitution effect was 1.110 kg which explains 1.1% of the phenotypic variance. Genes located closer to this SNP (±600 kb) included *LOC104976093* (uncharacterized LOC104976093) and *NRG1* (neuregulin 1) (Figure 6-Supplements). In this case, *NRG1* was the most important gene identified. This gene is located at 567.1 kb downstream of rs136155567. It is considered the direct ligand for ERBB3 and ERBB4 tyrosine kinase receptors. The multiple isoforms perform diverse functions such as inducing growth and differentiation of epithelial, glial, neuronal, and skeletal muscle cells, and influence the motor and sensory neuron development (Plowman *et al*., 1993; Ieguchi *et al*., 2010). In cattle, *NRG1* has been highly associated with organs development (Sweeney *et al*., 2001). Zhao *et al*. (2013) mentioned this gene as an emerging regulator of prolactin secretion, and it could influence weaning weight (Zhao *et al*., 2013).

The phenotypic variance explained by the SNPs identified in this study was very small (1.39 % on average). In growth trait studies, it is expected that most SNP markers will explain only a small proportion of the observed phenotypic variance, due to the polygenic control over such traits and because of individual genes influence phenotype only slightly. However, consideration of the set of SNPs that are significantly associated with each trait may allow a greater proportion of phenotypic variance to be explained. For example, the two SNP associated with WW explained 4.08 % of the variance in that trait.

CONCLUSION

The present GWAS identified three SNPs associated with growth traits of Braunvieh cattle. Two of them were in intergenic regions, the last one was located in an intronic region of *ARHGEF33* gene. However, there is evidence that some of the genes closer to the three identified SNPs markers are functionally related to growth. Four candidate genes were found to be potentially associated with growth traits in Braunvieh cattle, *MCM2*, *TPRA1*, GALM, and *NRG1*. Subsequent studies examining these genomic regions could lead to the identification of polymorphisms with potential uses in marker-assisted selection, providing a deeper understanding of the genetic basis of growth traits in cattle. Further analysis using the present information as a basis would allow conducting assessments on the ontogeny and specific search of causative mutations for live weight traits. Furthermore, examination of particular and general genic effects would indicate the possibility to include genomic information into current genetic evaluations.

ACKNOWLEDGMENTS

Authors thank the National Council for Science and Technology (CONACYT) for the financial support of the Doctorate studies of the first author and to the Universidad Autónoma Chapingo for the research grant DGIP-166701012. The authors thank the Asociación Mexicana de Criadores de Ganado Suizo de Registro, A. C. for their collaborative support.

LITERATURE CITED

Aki, T., T. Funakoshi, J. Nishida-Kitayama, and Y. Mizukami. 2008. TPRA40/GPR175 regulates early mouse embryogenesis through functional membrane transport by Sjögren's syndrome-associated protein NA14. J. Cell. Physiol. 217:194-206. doi: 10.1002/jcp.21492.

AMCGSR (Asociación Mexicana de Criadores de Ganado Suizo de Registro). 2017. Origen del ganado Pardo Suizo Europeo. (http://www.amcgsr.com.mx/del-suizo-europeo.php). Consulted on October 28th, 2017.

Aulchenko, Y. S., S. Ripke, A. Isaacs, and C. M. Van Duijn. 2007. GenABEL: an R library for genome-wide association analysis. Bioinformatics. 23:1294-1296. doi:10.1093/bioinformatics/btm108.

Beef Improvement Federation. 2016. Guidelines for uniform beef improvement programs. Ed. Jane Parish. 9th. Ed. Massachussets, U.S.A.

Ben Jemaa, S., M. Boussaha, M. Ben Mehdi, J. Heon Lee, and S. H. Lee. 2015. Genome-wide insights into population structure and genetic history of tunisian local cattle using the illumina bovinesnp50 beadchip. BMC Genomics. 16:677. doi: 10.1186/s12864-015-1638-6.

Boichard, D., C. Grohs, F. Bourgeois, F. Cerqueira, R. Faugeras, A. Neau, R. Rupp, Y. Amigues, M. Y. Boscher, and H. Levéziel. 2003. Detection of genes influencing economic traits in three French dairy cattle breeds. Genet. Sel. Evol. 35:77-101. doi: 10.1051/gse:2002037.

Buitenhuis, A. J., M. S. Lund, J. R. Thomasen, B. Thomsen, V. Hunnicke Nielsen, C. Bendixen, and B. Guldbrandtsen. 2007. Detection of quantitative trait loci affecting lameness and leg conformation traits in Danish Holstein cattle. J. Dairy Sci. 90:472-481. doi: 10.3168/jds.S0022-0302(07)72649-8.

Chin-Colli, R. C., R. Estrada-León, J. Magaña-Monforte, J. Segura-Correa, and R. Núñez-Domínguez. 2016. Genetic parameters for growth and reproductive traits of Brown Swiss cattle from Mexico. ERA. 3(7):11-20. doi: 10.19136/era.a3n7.167.

Elsik, C. G., D. R. Unni, C. M. Diesh, A. Tayal, M. L. Emery, H. N. Nguyen, and D. E. Hagen. 2016. Bovine genome database: New tools for gleaning function from the *Bos taurus* genome. Nucleic Acids Res. 4;44(D1):D834-9. doi:10.1093/nar/gkv1077.

Erbe, M., B. J. Hayes, L. K. Matukumalli, S. Goswam, P. J. Bowman, C. M. Reich, B. A. Mason, and M. E. Goddard. 2012. Improving accuracy of genomic predictions within and between dairy cattle breeds with imputed high-density single nucleotide polymorphism panels. J Dairy Sci. 95:4114–29. doi: 10.3168/jds.2011-5019.

Gao, J., Q. Wang, C. Dong, S. Chen, Y. Qi, and Y. Li. 2015. Whole exome sequencing identified mcm2 as a novel causative gene for autosomal dominant nonsyndromic deafness in a Chinese family. PLoS ONE 10(7):e0133522. doi:10.1371/journal.

Gondro, C., L. R. Porto-Neto, and S. H. Lee. 2014. SNPQC—an R pipeline for quality control of Illumina SNP genotyping array data. Anim. Genet. 45:758-761. doi: 10.1111/age.12198

Guo, J., H. Jorjain, and O. Carlborg. 2012. A genome-wide association study using international breeding-evaluation data identifies major loci affecting production traits and stature in the Brown Swiss cattle breed. BMC Genet. 13:82. doi: 10.1186/1471-2156-13-82.

Hagger, C., and A. Hofer. 1990. Genetic analyses of calving traits in the Swiss Black and White, Braunvieh and Simmental breeds by REML and MAPP procedures. Liv. Prod. Sci. 24:93-107. doi: 10.1016/0301-6226(90)90070-M.

Harris, B. L., and D. L. Johnson. 2010. Genomic predictions for New Zealand dairy bulls and integration with national genetic evaluation. J. Dairy Sci. 93:1243–52. doi: 10.3168/jds.2009-2619.

Hiersche, M., F. Rühle, and M. Stoll. 2013. Postgwas: Advanced GWAS Interpretation in R. PLoS ONE. 8:e71775. doi: 10.1371/journal.pone.0071775.

Hu, Z. L., C. A. Park, X. L. Wu, and J. M. Reecy. 2013. Animal QTLdb: an improved database tool for livestock animal QTL/association data dissemination in the post-genome era. Nucleic Acids Res. 41:871-879. doi: 10.1093/nar/gks1150.

Ieguchi, K., M. Fujita, Z. Ma, P. Davari, Y. Taniguchi, K., Sekiguchi, B. Wang, Y. K. Takada, and Y. Takada. 2010. Direct binding of the EGF-like domain of neuregulin-1 to integrins (αvβ3 and α6β4) is involved in Neuregulin-1/ErbB signaling. J. Biol. Chem. 285(41):31388-31398. doi:10.1074/jbc.M110.113878.

Jahuey-Martínez F. J., G. M. Parra-Bracamonte, A. M. Sifuentes-Rincón, J. C. Martínez-González, C. Gondro, C. A. García-Pérez, and L. A. López-Bustamante. 2016.

Genomewide association analysis of growth traits in Charolais beef cattle. J. Anim. Sci. 94:4570-4582. doi: 10.2527/jas.2016-0359.

Lu, D., S. Miller, M. Sargolzaei, M. Kelly, G. Vander Voort, T. Caldwell, Z. Wang, G. Plastow, and S. Moore. 2013. Genome-wide association analyses for growth and feed efficiency traits in beef cattle. J. Anim. Sci. 92:3612-3633. doi:10.2527/jas.2012-5716

Martínez, R., D. Bejarano, Y. Gómez, R. Dasoneville, A. Jiménez, G. Even, J. Sölkner, and G. Mészáros. 2016. Genome-wide association study for birth, weaning and yearling weight in Colombian Brahman cattle. Genet. Mol. Biol. 40(2): 453-459. doi: 10.1590/1678-4685-GMB-2016-0017.

Maxa, J., M. Neuditschko, I. Russ, M. Förster, and I. Medugorac. 2012. Genome-wide association mapping of milk production traits in Braunvieh cattle. J. Dairy Sci. 95:5357-5364. doi: 10.3168/jds.2011-4673.

McClure, M. C., N. S. Morsci, R. D. Schnabel, J. W. Kim, P. Yao, M. M. Rolf, S. D. McKay, S. J. Gregg, R. H. Chapple, S. L. Northcutt, and J. F. Taylor. 2010. A genome scan for quantitative trait loci influencing carcass, post-natal growth and reproductive traits in commercial Angus cattle. Anim. Genet. 41:597–607. doi:10.1111/j.1365-2052.2010.02063.x.

Orantes-Zebadúa, M. A., D. Platas-Rosado, V. Córdova-Avalos, M. C. Santos-Lara, and A. Córdova-Avalos. 2014. Characterization of dual purpose livestock in a Region of Chiapas, Mexico. Ecosistemas y Recursos Agropecuarios. 1(1):49-58. doi: 10.19136/era.a1n1.6.

Phillips, W. A., J. W. Holloway, B. Warrington, and B. C. Venuto. 2009. Case study: Stocker and feedlot performance of beef heifers sired by Braunvieh and Wagyu

Bulls from Angus, Brahman, Senepol, and Tuli-sired Dams. The Professional Animal Scientist. 25:809-814.

Plieschke L., C. Edel, E. C. G. Pimentel, R. Emmerling, J. Bennewitz, and K. U. Götz. 2015. A simple method to separate base population and segregation effects in genomic relationship matrices. Genet. Sel. Evol. 47:53. doi: 10.1186/s12711-015-0130-8.

Plowman, G. D., J. M. Green, J. M. Culouscou, G. W. Carlton, V. M. Rothwell, and S. Buckley. 1993. Heregulin induces tyrosine phosphorylation of HER4/p180^{erb4}. Nature 366(6454):473-475. doi: 10.1038/366473a0.

Price, A. L., N. J. Patterson, R. M. Plenge, M. E. Weinblatt, N. A. Shadick, and D. Reich. 2006. Principal components analysis corrects for stratification in genome-wide association studies. Nat. Genet. 38:904-909. doi:10.1038/ng1847.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945–959.

Purfield, D. C., D. G. Bradley, R. D. Evans, F. J. Kearney, and D. P. Berry. 2015. Genome-wide association study for calving performance using high-density genotypes in dairy and beef cattle. Gen. Sel. Evol. 47:47. doi: 10.1186/s12711-015-0126-4.

Rolf, M. M., J. F. Taylor, R. D. Schnabel, S. D. McKay, M. C. McClure, S. L. Northcutt, M. S. Kerley, and R. L. Weaber. 2012. Genome-wide association analysis for feed efficiency in Angus cattle. Anim. Genet. 43:367-374. doi: 10.1111/j.1365-2052.2011.02273.x.

Sherman, E. L., J. D. Nkrumah, C. Li, R. Bartusiak, B. Murdoch, and S. S. Moore. 2009. Fine mapping quantitative trait loci for feed intake and feed efficiency in beef cattle. J. Anim. Sci. 87:37-45. doi: 10.2527/jas.2008-0876.

Shin, D. H., H. J. Lee, S. Cho, H. J. Kim, J. Y. Hwang, C. K. Lee, J. Y. Jeong, D. Yoon, and H. Kim. 2014. Deleted copy number variation of Hanwoo and Holstein using next generation sequencing at the population level. BMC Genomics 15:240. doi: 10.1186/1471-2164-15-240.

Silva, C., R. Aké, and R. Valle. 2002. Edad y crecimiento a la pubertad en toros Suizo Pardo en condiciones tropicales. Cuban J. Agr. Sci. 36(3):205-210.

Singh, J., X. Wen, and S. J. Scales. 2015. The orphan G protein-coupled receptor Gpr175 (Tpra40) enhances hedgehog signaling by modulating cAMP levels. J. Biol. Chem. 290(49):29663-29675. doi: 10.1074/jbc.M115.665810.

Smitz, N., D. Cornélis, P. Chardonnet, A. Caron, M. Garine-Wichatitsky, F. Jori, A. Mouton, A. Latinne, L. M. Pigneur, M. Melletti, K. L Knapeckas, J. Marescaux, C. L. Pereira, and J. Michaux. 2014. Genetic structure of fragmented southern populations of African Cape buffalo (*Syncerus caffer caffer*). BMC Evol. Biol. 14:203. doi: 10.1186/s12862-014-0203-2.

Sweeney, C., D. Fambrough, C. Huard, A. J. Diamonti, E. S. Lander, L. C. Cantley, and K. L. Carraway. 2001. Growth factor-specific signaling pathway stimulation and gene expression mediated by ErbB receptors. J. Biol. Chem. 275(25):22685-22698. doi:10.1074/jbc.M100602200.

Takasuga A. 2016. PLAG1 and NCAPG-LCORL in livestock. Anim. Sci. J. 87(2):159-167. doi: 10.1111/asj.12417.

Thoden, J. B., D. J. Timson, R. J. Reece, and H. M. Holden. 2004. Molecular structure of human galactose mutarotase. J. Biol. Chem. 279(22):23431-23437. doi:10.1074/jbc.M40234720.

Todorov, I. T., R. Pepperkok, R. N. Philipova, S. E. Kearsey, W. Ansorge, and D. Werner. 1994. A human nuclear protein with sequence homology to a family of early S phase proteins is required for entry into S phase and for cell division. J. Cell Sci. 107:253-265.

UniProt: the universal protein knowledgebase The UniProt Consortium. 2017. Nucleic Acids Research, 45(D1): D158–D169.

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

Zhao, W. 2013. Neuregulin-1 (Nrg1): An emerging regulator of prolactin (PRL) secretion. *In:* "Prolactin". Ed. György M. Nagy. IntechOpen, London. United Kingdom. doi:10.5772/54716.

Zhang H., Z. Wang, S. Wang, and H. Li. 2012. Progress of genome wide association study in domestic animals. J. Anim. Sci. Biotech. 3:26. doi: 10.1186/2049-1891-3-26.

Zimin, A. V., A. L. Delcher, L. Florea, D. R. Kelley, M. C. Schatz, D. Puiu, F. Hanrahan, G. Pertea, C. V. Van Tassel, T. S. Sonstegard, G. Marçais, M. Roberts, P. Subramanian, J. A. Yprke, and S. L. Salzberg. 2009. A whole-genome assembly of the domestic cow, *Bos taurus*. Genome Biol. 10:R42. doi: 10.1186/gb-2009-10-4-r42Tables.

SUPPLEMENTS

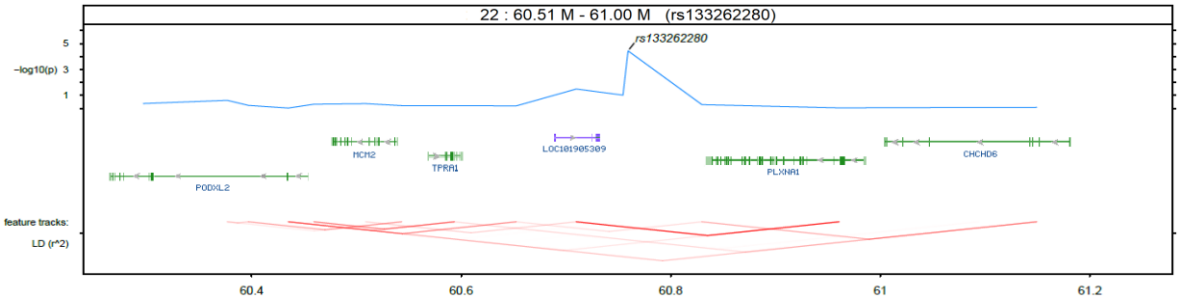

Figure 4. Genes located close to rs133262280 marker and linkage disequilibrium (LD) of the markers present in the represented region of the BTA22.

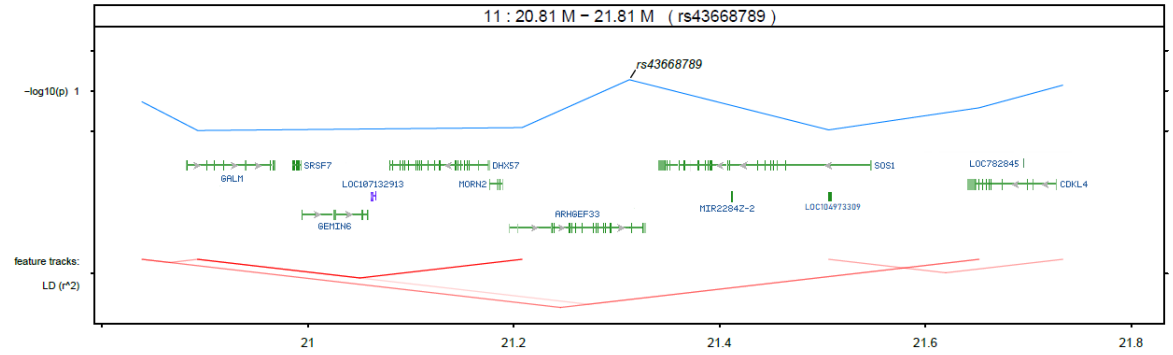

Figure 5. Genes located close to rs43668789 marker and linkage disequilibrium (LD) of the markers present in the represented region of the BTA 11.

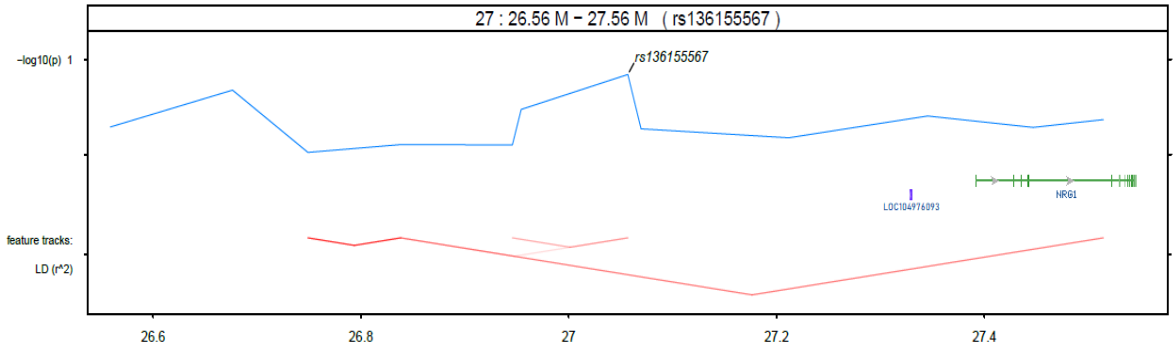

Figure 6. Genes located close to rs136155567 marker and linkage disequilibrium (LD) of the markers present in the represented region of the BTA 27.