POLIMORFISMOS EN GÉNES 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………………………………………………………………………… vii |
| LISTA DE FIGURAS………………………………………………………………………… ix |
| DEDICATORIAS……………………………………………………………………………… xi |
| AGRADECIMIENTOS……………………………………………………………………… xii |
| DATOS BIOGRÁFICOS……………………………………………………………………… xiii |
| RESUMEN………………………………………………………………………………………… 1 |
| ABSTRACT………………………………………………………………………………………… 2 |
| 1. INTRODUCCIÓN GENERAL…………………………………………………………… 3 |
| 1.1. Hipótesis………………………………………………………………………………… 4 |
| 1.2. Objetivos………………………………………………………………………………… 5 |
| 1.3. Estructura de la tesis………………………………………………………………… 5 |
| 1.4. Literatura citada……………………………………………………………………….. 6 |
| 2. SOURCES OF BIAS IN GENETIC ASSOCIATION STUDIES OF CATTLE: A REVIEW………………………………………………………………………………… 8 |
| 2.1. Abstract………………………………………………………………………………….. 10 |
| 2.2. Resumen………………………………………………………………………………….. 10 |
| 2.3. Resumo………………………………………………………………………………….. 11 |
| 2.4. Introduction……………………………………………………………………………… 12 |
| 2.5. Source of bias in genetic association studies………………………………………. 13 |
| 2.5.1. Genetic factors…………………………………………………………………………… 14 |
| 2.5.2. Genomic factors…………………………………………………………………………… 14 |
| 2.5.3. Sampling-related factors……………………………………………………………… 17 |
| 2.5.4. Statistical factors……………………………………………………………………….. 19 |
| 2.6. Strategies to correct biases in GAS…………………………………………………… 21 |
| 2.4.1. Genetic-genomic strategies…………………………………………………………… 22 |
| 2.4.2. Sampling-related and statistical strategies……………………………………….. 23 |
| 2.7. Acknowledgments……………………………………………………………………….. 24 |
2.8. Conflicts of interest................................................................. 24
2.9. References........................................................................... 25

3. POTENTIAL INFLUENCE OF CSN3 AND LGB GENES IN GENETIC
ASSOCIATION STUDIES OF MILK QUALITY TRAITS...................... 35
3.1. Abstract................................................................................ 37
3.2. Introduction........................................................................... 38
3.3. Materials and methods.......................................................... 38
   3.3.1. Paper selection criteria.................................................. 38
   3.3.2. Meta-analysis................................................................. 39
3.4. Results and discussion............................................................ 40
3.5. Conclusion........................................................................... 43
3.6. Conflict of interest................................................................. 43
3.7. Acknowledgments................................................................. 43
3.8. References........................................................................... 44

4. POTENTIAL CANDIDATE GENES FOR INTEGRAL GENETIC
IMPROVEMENT PROGRAMS IN BEEF CATTLE......................... 51
4.1. Abstract................................................................................ 53
4.2. Introduction........................................................................... 54
4.3. Genetic association studies in beef cattle ......................... 55
4.4. Candidate genes associated with economical important traits in beef
cattle......................................................................................... 56
4.5. Candidate genes consortiums............................................... 59
   4.5.1. CAPN-CAST............................................................... 60
   4.5.2. CHCHD7-PLAG1...................................................... 60
   4.5.3. FABP4-PPARγ.......................................................... 61
   4.5.4. NCAPG-LCORL.......................................................... 61
4.6. Genetic architecture and identification of consortium genes with
pleiotropic effect........................................................................ 62
4.7. Potential use of pleiotropic and candidate gene consortiums...... 64
4.8. Conclusions and implications.............................................. 66
4.9. Acknowledgments................................................................. 66
<table>
<thead>
<tr>
<th>Cuadro</th>
<th>Capítulo</th>
<th>Título</th>
<th>Página</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>Results of genetic association studies between CSN3 gene with milk yield in dairy cattle</td>
<td>12</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>Former and current approaches used in genetic association studies</td>
<td>13</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>Level of significance for the effects included in the models to analyze milk production (MP), and protein (PP) and fat (FP) percentage in milk</td>
<td>47</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>Least square means and multiple comparison for milk production for the interaction Breed by Gene</td>
<td>48</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>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 and LGB gene</td>
<td>49</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>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</td>
<td>50</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>Candidate genes associated with two groups of traits usually considered in genetic improvement programs</td>
<td>77</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>Candidate genes associated with two groups of traits usually considered in genetic improvement programs (II)</td>
<td>78</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>Candidate genes associated with three groups of traits usually considered in genetic improvement programs</td>
<td>79</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>Candidate genes associated with four groups of traits usually considered in genetic improvement programs</td>
<td>80</td>
</tr>
<tr>
<td>Cuadro</td>
<td>Capítulo</td>
<td>Título</td>
<td>Página</td>
</tr>
<tr>
<td>--------</td>
<td>----------</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>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</td>
<td>95</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>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)</td>
<td>96</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>Allele and genotype frequency of markers previously associated to genetic diseases present in the studied population</td>
<td>97</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>Allele and genotype frequency of markers previously associated to genetic diseases present in the studied population (Continue I)</td>
<td>98</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>Allele and genotype frequency of markers previously associated to genetic diseases present in the studied population (Continue II)</td>
<td>99</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>Allele and genotype frequency of markers previously associated to genetic diseases present in the studied population (Continue III)</td>
<td>100</td>
</tr>
<tr>
<td>1</td>
<td>6</td>
<td>Descriptive statistics for growth traits of Braunvieh Cattle</td>
<td>106</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>Parameters and statistics of SNP associated with growth traits of Braunvieh cattle</td>
<td>112</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>Genes and previously reported QTL located near SNP associated with growth traits of Braunvieh cattle</td>
<td>114</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>Genes close to the SNP associated with growth traits of Braunvieh cattle</td>
<td>115</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>Genes close to the SNP associated with growth traits of Braunvieh cattle. Continue</td>
<td>116</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>Previously reported QTL found near the SNP associated with growth traits of Braunvieh cattle</td>
<td>118</td>
</tr>
<tr>
<td>Figura</td>
<td>Capítulo</td>
<td>Título</td>
<td>Página</td>
</tr>
<tr>
<td>-------</td>
<td>----------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>Diversity and distribution of major <em>Bos taurus</em> and <em>Bos indicus</em> haplogroups</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>Information content variability for haplotype level in Europeans</td>
<td>16</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>Types of MAF according its effect size</td>
<td>17</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>Network of candidate pleiotropic genes for carcass traits in Nellore cattle</td>
<td>18</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>Sample sizes used in genetic association studies showing errors of type I</td>
<td>19</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>Forest plot of the effects or recombinant bovine somatotropin on the risk ratio of clinical mastitis</td>
<td>21</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>Probabilities of association obtained with two different Bayesian-based methods</td>
<td>22</td>
</tr>
<tr>
<td>1</td>
<td>6</td>
<td>Presence of two subpopulations in the sample of the Braunvieh population analyzed. Circles of different color indicate different population</td>
<td>110</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>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</td>
<td>111</td>
</tr>
<tr>
<td>Figura</td>
<td>Capítulo</td>
<td>Título</td>
<td>Página</td>
</tr>
<tr>
<td>--------</td>
<td>----------</td>
<td>------------------------------------------------------------------------</td>
<td>--------</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>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 &lt; 5 \times 10^{-5} )). BWT = birth weight; WW = weaning weight; YW = yearling weight</td>
<td>113</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>Genes located closer to rs133262280 marker and linkage disequilibrium (LD) of the markers present in the represented region of the BTA 22</td>
<td>128</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>Genes located closer to rs43668789 marker and linkage disequilibrium (LD) of the markers present in the represented region of the BTA 11</td>
<td>128</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>Genes located closer to rs136155567 marker and linkage disequilibrium (LD) of the markers present in the represented region of the BTA 27</td>
<td>128</td>
</tr>
</tbody>
</table>
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 (Ruiz-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 está 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


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

José Luis Zepeda-Batista¹, María Isabel Carrillo-Díaz², Luis Antonio Saavedra-Jiménez¹

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

²Facultad 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 Batista¹, MVZ, MSc; María I Carrillo Díaz²*, IAZ, PhD; Luis A Saavedra Jiménez¹, IAZ, MSc.

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

²Facultad 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.

<table>
<thead>
<tr>
<th>Author</th>
<th>Gustavsson et al. (2014)</th>
<th>Duifhuis-Rivera et al. (2014)</th>
<th>Deb et al. (2014)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sampled animals</td>
<td>400</td>
<td>202</td>
<td>200</td>
</tr>
<tr>
<td>Reported effect</td>
<td>Positive</td>
<td>Absence</td>
<td>Positive</td>
</tr>
<tr>
<td>Best genotype*</td>
<td>AA</td>
<td>N/D</td>
<td>AB</td>
</tr>
</tbody>
</table>

*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

<table>
<thead>
<tr>
<th>Approach</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Family Based</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TDT¹</td>
<td>Quality control; robustness to population stratification; ability to perform genotyping quality control</td>
<td>Less power than pop-based GWAS; computationally demanding; not practical for late-onset diseases</td>
</tr>
<tr>
<td>FB-GWAS²</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Candidate polymorphism &amp;</td>
<td>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</td>
<td>SNPs may not serve as the true trait-causing variants; multiple SNPs measurements are needed to know a precise location on the genome</td>
</tr>
<tr>
<td>Candidate gene</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Population based</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fine mapping</td>
<td>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</td>
<td>Data preprocessing extensive and computational burden greater; specific software requirements; need for candidate gene studies to validate findings from GWAS</td>
</tr>
<tr>
<td>Genome-wide</td>
<td>Identify associations between SNPs and a trait; involves the characterization of larger number of SNPs</td>
<td></td>
</tr>
</tbody>
</table>

¹TDT: transmission disequilibrium test; ²FB-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).](image)

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.
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).](image)

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 (Groenwald 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


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


http://cshprotocols.cshlp.org/cgi/pmidlookup?view=long&pmid=22383645


http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0158165

http://www.journalofdairyscience.org/article/S0022-0302(15)00084-3/pdf
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3480678/


http://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1000028


https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3468889/


https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5125095/

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3518898/

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.


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; \( A_i = \) random effect of \( i \)-th paper included in the study; \( B_z = \) \( z \)-th effect of the breed; \( M_y = \) \( y \)-th effect of the used model in the original study; \( T_x = \) \( x \)-th effect of the used test to recover the original data; \( A_G w = \) \( w \)-th effect of the analyzed gene; \( G_N V = \) \( v \)-th effect of the genotype nested in the analyzed gene; \( S S u = \) \( u \)-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 (\( A_G \)), CSN3 and LGB; genotype (\( G_N \)), AA, AB, AE, BB, BE, and BE for CSN3, and for LGB were AA, AB, and BB; sample size (\( S S \)), 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 (\( B A G \)); breed by genotype nested in analyzed gene (\( B G N \)); and model by genotype nested in analyzed gene (\( M G N \)). 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≤0.0255) 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


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

<table>
<thead>
<tr>
<th>Variable</th>
<th>B(^b)</th>
<th>M(^\upgamma)</th>
<th>T(^\alpha)</th>
<th>AG(^\omega)</th>
<th>GN(^\nu)</th>
<th>SS(^\eta)</th>
<th>BAG(^i)</th>
<th>BGN(^s)</th>
<th>MGN(^r)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MP</td>
<td>0.1528</td>
<td>0.2411</td>
<td>&lt;0.0001</td>
<td>0.1833</td>
<td>0.0974</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>0.0255</td>
</tr>
<tr>
<td>PP</td>
<td>0.7529</td>
<td>0.7358</td>
<td>0.6484</td>
<td>0.9966</td>
<td>1.000</td>
<td>0.9090</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FP</td>
<td>0.1512</td>
<td>0.4161</td>
<td>0.8409</td>
<td>0.8604</td>
<td>1.000</td>
<td>0.8199</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B\(^b\), breed; M\(^\upgamma\), model used in the original study; T\(^\alpha\), test used to recover original data; AG\(^\omega\), analyzed gene; GN\(^\nu\), genotype nested in analyzed gene; SS\(^\eta\), sample size; BAG\(^i\), 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

<table>
<thead>
<tr>
<th>Breed</th>
<th>CSN3</th>
<th>LGB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brown Swiss</td>
<td>5.60±0.46&lt;sup&gt;ab&lt;/sup&gt;</td>
<td>5.64±0.46&lt;sup&gt;ab&lt;/sup&gt;</td>
</tr>
<tr>
<td>Holstein</td>
<td>6.01±0.19&lt;sup&gt;a&lt;/sup&gt;</td>
<td>6.02±0.19&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Jersey</td>
<td>5.55±0.25&lt;sup&gt;b&lt;/sup&gt;</td>
<td>5.56±0.25&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Fleckvieh</td>
<td>5.98±0.22&lt;sup&gt;ab&lt;/sup&gt;</td>
<td>5.70±0.22&lt;sup&gt;ab&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a,b</sup> 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

<table>
<thead>
<tr>
<th>Breed</th>
<th>Genotype of the CSN3 gene</th>
<th>Genotype of the LGB gene</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AA</td>
<td>AB</td>
</tr>
<tr>
<td>Brown Swiss</td>
<td>5.54±0.47\textsuperscript{a}</td>
<td>5.62±0.47\textsuperscript{a}</td>
</tr>
<tr>
<td>Holstein</td>
<td>6.04±0.22\textsuperscript{a}</td>
<td>6.01±0.22\textsuperscript{a}</td>
</tr>
<tr>
<td>Jersey</td>
<td>5.59±0.31\textsuperscript{a}</td>
<td>5.57±0.31\textsuperscript{a}</td>
</tr>
<tr>
<td>Fleckvieh</td>
<td>5.80±0.24\textsuperscript{a}</td>
<td>5.86±0.24\textsuperscript{a}</td>
</tr>
</tbody>
</table>

\textsuperscript{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

<table>
<thead>
<tr>
<th>Model</th>
<th>Genotype of the CSN3 gene</th>
<th>Genotype of the LGB gene</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AA</td>
<td>AB</td>
</tr>
<tr>
<td>LS</td>
<td>5.96±0.18&lt;sup&gt;a&lt;/sup&gt;</td>
<td>5.98±0.18&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>MM</td>
<td>5.53±0.19&lt;sup&gt;b&lt;/sup&gt;</td>
<td>5.54±0.19&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>1</sup>LS, least square model; MM, mixed model based on bayesian and ridge regression.

<sup>a</sup><sup>b</sup><sup> </sup>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-Batista¹, Gaspar Manuel Parra Bracamonte², Rafael Núñez Domínguez¹, Rodolfo Ramírez Valverde¹, Agustín Ruíz Flores¹


Submitted to the Italian Journal of Animal Science
Potential candidate genes for integral genetic improvement programs in
beef cattle

José Luis Zepeda Batista¹, Gaspar Manuel Parra Bracamonte², Rafael Núñez
Domínguez¹€, Rodolfo Ramírez Valverde¹, Agustín Ruíz Flores¹
¹Universidad Autónoma Chapingo. Departamento de Zootecnia. Posgrado en
Producción Animal. Km. 38.5 Carretera México-Texcoco. Chapingo, Estado de
México. C.P. 56230
²Laboratorio de Biotecnología Animal. Centro de Biotecnología Genómica.
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 establish 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 promotior 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 promoters 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-traitst 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


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


37. Dujková R, Ranganathan Y, Dufek A, Macák J, Bezdícek J. Polymorphic effects of FABP4 and SCD genes on intramuscular fatty acid profiles in


44. Sasago N, Abe T, Sakuma H, Kojima T, Uemoto Y. Genome-wide association study for carcass traits, fatty acid composition, chemical


## Tables

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

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

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

BT: backfat thickness; BW: 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

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

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

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

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-Batista¹, G.M. Parra-Bracamonte², R. Núñez-Domínguez¹, R. Ramírez-Valverde¹, A. Ruíz-Flores¹


Submitted to the Tropical Animal Health and Production
Screening genetic diseases prevalence in Braunvieh cattle

Genetic diseases prevalence in Braunvieh cattle

José Luis Zepeda-Batista¹: jzepedab@gmail.com
Gaspar Manuel Parra-Bracamontes²: gparra@ipn.mx
Rafael Núñez-Domínguez¹: rafael.nunez@correo.chapingo.mx
Rodolfo Ramírez-Valverde¹: rodolfov@correo.chapingo.mx
Agustín Ruíz-Flores¹: arf@correo.chapingo.mx


Corresponding author:
Rafael Núñez-Domínguez¹: 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


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


### 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.

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

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

<sup>a</sup>Fq Ht: frequency of heterozygotes. <sup>b</sup>Fq Ho: frequency of homozygotes. N/P: animals with disease associated markers not present in the population. <sup>c</sup>Deficiency of uridine monophosphate synthase. <sup>d</sup>Syndactyly syndrome; <sup>e</sup>Bovine leukocyte adhesion deficiency; <sup>f</sup>Hereditary perinatal weak calf syndrome; <sup>g</sup>Bovine progressive degenerative myeloencephalopathy.
Table 2. Allele and genotype frequency of markers previously associated with genetic diseases present in the studied population.

<table>
<thead>
<tr>
<th>Disease</th>
<th>Gene Symbol</th>
<th>Marker Name</th>
<th>Marker position</th>
<th>Chr</th>
<th>SNP*</th>
<th>Allele*</th>
<th>Genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1 2</td>
<td>11 12 22</td>
</tr>
<tr>
<td>BMS^a</td>
<td>FBN1</td>
<td>FBN1_1</td>
<td>62054844</td>
<td>10</td>
<td>G/A</td>
<td>0.998 0.002</td>
<td>0.997 0.003</td>
</tr>
<tr>
<td>BSD^b</td>
<td>SPAST</td>
<td>SDM</td>
<td>n/e</td>
<td>11</td>
<td>G/A</td>
<td>0.998 0.002</td>
<td>0.997 0.003</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SDM_2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SDM_3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BH2^c</td>
<td>TUBD1</td>
<td>BH2</td>
<td>n/e</td>
<td>24</td>
<td>T/C</td>
<td>0.997 0.003</td>
<td>0.997 n/p</td>
</tr>
<tr>
<td>Citrullinaemia</td>
<td>ASS1</td>
<td>ASS1</td>
<td>100802781</td>
<td>11</td>
<td>C/T</td>
<td>0.997 0.003</td>
<td>0.997 n/p</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Citrullinemia_3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CMD^d</td>
<td>ATP2A1</td>
<td>CMD1</td>
<td>n/e</td>
<td>25</td>
<td>T/C</td>
<td>0.997 0.003</td>
<td>0.997 n/p</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CMD1_3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CMD2_2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crooked Tail</td>
<td>MRC2</td>
<td>MRC2_1</td>
<td>47734925</td>
<td>19</td>
<td>I/D</td>
<td>0.995 0.005</td>
<td>0.990 0.010</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CTS-BB</td>
<td>47740444</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CTS-BB_2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CTS-BB_3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DUMPS^e</td>
<td>UMPS</td>
<td>DUMPS</td>
<td>n/e</td>
<td>1</td>
<td>C/T</td>
<td>0.998 0.002</td>
<td>0.997 0.003</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DUMPS_2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>DUMPS_3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disease</td>
<td>Gene Symbol</td>
<td>Marker Name</td>
<td>Marker position</td>
<td>Chr</td>
<td>SNP*</td>
<td>Allele*</td>
<td>Genotype</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
<td>-------------------</td>
<td>-----------------</td>
<td>-----</td>
<td>------</td>
<td>---------</td>
<td>----------</td>
</tr>
<tr>
<td>EB</td>
<td>LAMC2</td>
<td>EB_2</td>
<td>4164**</td>
<td>5</td>
<td>G/A</td>
<td>0.998</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>695072</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypotrichosis</td>
<td>HEPHL1</td>
<td>HEPHL1</td>
<td>29</td>
<td>29</td>
<td>A/T</td>
<td>0.998</td>
<td>0.002</td>
</tr>
<tr>
<td>Mulefoot</td>
<td>LRP4</td>
<td>LRP4_3</td>
<td>77675440</td>
<td>15</td>
<td>C/A</td>
<td>0.705</td>
<td>0.295</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mulefoot-241_2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mulefoot-241_3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mulefoot-2719</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>MSUD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>MSUD_2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>MSUD_3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSUD</td>
<td>BCKDHA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMA</td>
<td>AFG3L2</td>
<td>SMA</td>
<td></td>
<td>24</td>
<td>G/T</td>
<td>0.997</td>
<td>0.003</td>
</tr>
<tr>
<td>Pompes</td>
<td>GAA</td>
<td>Pompes_1783_BR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pompes_1783_BR_2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pompes_1783_BR_3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WCS</td>
<td>IARS</td>
<td>IARS</td>
<td>85341291</td>
<td>8</td>
<td>G/C</td>
<td>0.998</td>
<td>0.002</td>
</tr>
</tbody>
</table>
Table 2. Allele and genotype frequency of markers previously associated with genetic diseases present in the studied population (Continue II).

<table>
<thead>
<tr>
<th>Disease</th>
<th>Gene Symbol</th>
<th>Marker Name</th>
<th>Marker position</th>
<th>Chr</th>
<th>SNP*</th>
<th>Allele* 1</th>
<th>Allele* 2</th>
<th>Genotype 11</th>
<th>Genotype 12</th>
<th>Genotype 22</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>WVR_49656945</td>
<td>49656945</td>
<td>C/T</td>
<td>0.996</td>
<td>0.004</td>
<td>0.993</td>
<td>0.007</td>
<td>n/p</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WVR_49657798</td>
<td>49657798</td>
<td>G/A</td>
<td>0.989</td>
<td>0.011</td>
<td>0.986</td>
<td>0.007</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WVR_49664852</td>
<td>49664852</td>
<td>C/T</td>
<td>0.996</td>
<td>0.004</td>
<td>0.993</td>
<td>0.007</td>
<td>n/p</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WVR_49667361</td>
<td>49667361</td>
<td>C/T</td>
<td>0.996</td>
<td>0.004</td>
<td>0.993</td>
<td>0.007</td>
<td>n/p</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WVR_49673503</td>
<td>49673503</td>
<td>C/T</td>
<td>0.993</td>
<td>0.007</td>
<td>0.987</td>
<td>0.013</td>
<td>n/p</td>
<td></td>
</tr>
<tr>
<td>WS</td>
<td>PNPLA8</td>
<td>WVR_49681169</td>
<td>49681169</td>
<td>C/T</td>
<td>0.996</td>
<td>0.004</td>
<td>0.993</td>
<td>0.007</td>
<td>n/p</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WVR_49682552</td>
<td>49682552</td>
<td>A/G</td>
<td>0.996</td>
<td>0.004</td>
<td>0.993</td>
<td>0.007</td>
<td>n/p</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WVR_49686038</td>
<td>49686038</td>
<td>A/G</td>
<td>0.996</td>
<td>0.004</td>
<td>0.993</td>
<td>0.007</td>
<td>n/p</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WVR_49687224</td>
<td>49687224</td>
<td>T/G</td>
<td>0.996</td>
<td>0.004</td>
<td>0.993</td>
<td>0.007</td>
<td>n/p</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WVR_49691015</td>
<td>49691015</td>
<td>G/T</td>
<td>0.989</td>
<td>0.011</td>
<td>0.986</td>
<td>0.007</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WVR_49692015</td>
<td>49692015</td>
<td>T/C</td>
<td>0.996</td>
<td>0.004</td>
<td>0.993</td>
<td>0.007</td>
<td>n/p</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WVR_49692485</td>
<td>49692485</td>
<td>G/T</td>
<td>0.989</td>
<td>0.011</td>
<td>0.986</td>
<td>0.007</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WVR_49692825</td>
<td>49692825</td>
<td>C/T</td>
<td>0.990</td>
<td>0.010</td>
<td>0.980</td>
<td>0.020</td>
<td>n/p</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WVR_49693121</td>
<td>49693121</td>
<td>C/T</td>
<td>0.996</td>
<td>0.004</td>
<td>0.993</td>
<td>0.007</td>
<td>n/p</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WVR_49693140</td>
<td>49693140</td>
<td>G/T</td>
<td>0.993</td>
<td>0.007</td>
<td>0.987</td>
<td>0.013</td>
<td>n/p</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WVR_49693164</td>
<td>49693164</td>
<td>A/T</td>
<td>0.996</td>
<td>0.004</td>
<td>0.993</td>
<td>0.007</td>
<td>n/p</td>
<td></td>
</tr>
<tr>
<td>Disease</td>
<td>Gene Symbol</td>
<td>Marker Name</td>
<td>Marker position</td>
<td>Chr</td>
<td>SNP</td>
<td>Allele 1</td>
<td>Allele 2</td>
<td>Genotype 11</td>
<td>Genotype 12</td>
<td>Genotype 22</td>
</tr>
<tr>
<td>-------------</td>
<td>-------------</td>
<td>---------------</td>
<td>-----------------</td>
<td>-----</td>
<td>--------</td>
<td>----------</td>
<td>----------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------</td>
</tr>
<tr>
<td>WS</td>
<td>PNPLA8</td>
<td>WVR_49693265</td>
<td>49693265</td>
<td>A/G</td>
<td>0.996</td>
<td>0.004</td>
<td>0.993</td>
<td>0.007</td>
<td>n/p</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WVR_49693601</td>
<td>49693601</td>
<td>T/C</td>
<td>0.996</td>
<td>0.004</td>
<td>0.993</td>
<td>0.007</td>
<td>n/p</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WVR_49693913</td>
<td>49693913</td>
<td>A/C</td>
<td>0.996</td>
<td>0.004</td>
<td>0.993</td>
<td>0.007</td>
<td>n/p</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WVR_49694977</td>
<td>49694977</td>
<td>A/G</td>
<td>0.996</td>
<td>0.004</td>
<td>0.993</td>
<td>0.007</td>
<td>n/p</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WVR_49695504</td>
<td>49695504</td>
<td>A/G</td>
<td>0.993</td>
<td>0.007</td>
<td>0.987</td>
<td>0.013</td>
<td>n/p</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WVR_49695952</td>
<td>49695952</td>
<td>C/T</td>
<td>0.989</td>
<td>0.011</td>
<td>0.986</td>
<td>0.007</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WVR_49698002</td>
<td>49698002</td>
<td>C/T</td>
<td>0.993</td>
<td>0.007</td>
<td>0.987</td>
<td>0.013</td>
<td>n/p</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WVR_49698436</td>
<td>49698436</td>
<td>C/T</td>
<td>0.996</td>
<td>0.004</td>
<td>0.993</td>
<td>0.007</td>
<td>n/p</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WVR_49700154</td>
<td>49700154</td>
<td>A/G</td>
<td>0.996</td>
<td>0.004</td>
<td>0.993</td>
<td>0.007</td>
<td>n/p</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WVR_49701106</td>
<td>49701106</td>
<td>T/C</td>
<td>0.993</td>
<td>0.007</td>
<td>0.987</td>
<td>0.013</td>
<td>n/p</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WVR_49702287</td>
<td>49702287</td>
<td>T/C</td>
<td>0.996</td>
<td>0.004</td>
<td>0.993</td>
<td>0.007</td>
<td>n/p</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WVR_49702494</td>
<td>49702494</td>
<td>G/A</td>
<td>0.996</td>
<td>0.004</td>
<td>0.993</td>
<td>0.007</td>
<td>n/p</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WVR_49708283</td>
<td>49708283</td>
<td>T/C</td>
<td>0.996</td>
<td>0.004</td>
<td>0.993</td>
<td>0.007</td>
<td>n/p</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WVR_49715678</td>
<td>49715678</td>
<td>C/T</td>
<td>0.996</td>
<td>0.004</td>
<td>0.993</td>
<td>0.007</td>
<td>n/p</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WVR_49718641</td>
<td>49718641</td>
<td>C/T</td>
<td>0.996</td>
<td>0.004</td>
<td>0.993</td>
<td>0.007</td>
<td>n/p</td>
<td></td>
</tr>
</tbody>
</table>

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


Submitted to the Journal of Animal Science
Running head: GWAS for growth traits in Mexican Braunvieh cattle

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


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

<table>
<thead>
<tr>
<th>Trait</th>
<th>N</th>
<th>n(QC)²</th>
<th>Mean</th>
<th>SD</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>BWT</td>
<td>300</td>
<td>266</td>
<td>38.01</td>
<td>4.07</td>
<td>22</td>
<td>50</td>
</tr>
<tr>
<td>WW</td>
<td>300</td>
<td>263</td>
<td>212.40</td>
<td>27.43</td>
<td>128</td>
<td>308</td>
</tr>
<tr>
<td>YW</td>
<td>300</td>
<td>244</td>
<td>313.17</td>
<td>45.47</td>
<td>176</td>
<td>440</td>
</tr>
</tbody>
</table>

¹BWT: birth weight; WW: weaning weight; YW = yearling weight.
²n(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 \((\beta)\) 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](image)

**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. BW = birth weight; WW = weaning weight; YW = yearling weight.

According to the significance threshold considered (P < 5 × 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

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

1BWT = birth weight; WW = weaning weight.
2ID = identification.
3UMD version 3.1 (Zimin et al., 2009).
4Elsik et al. (2015).
5MAF = 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.
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

---

**Table 3. Genes and previously reported QTL1 located near significant associated SNP**

<table>
<thead>
<tr>
<th>Trait_SNP ID2_BTA_Mb</th>
<th>Genes in ±250 kb</th>
<th>QTL</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>BW_T_rs133262280_22_60.7</td>
<td>PODXL2, MCM2, TPRA1, LOC10105309, PLXNA1, CHCHD6</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>WW_rs43668789_11_21.3</td>
<td>ARHGEF33, MORN2, DHX57, GEMIN6, SRSF7, GALM, MIR2284Z-2, SOS1, CDKL4, MAP4K3</td>
<td>SOUND</td>
<td>Buitenhuis et al., 2007</td>
</tr>
<tr>
<td>WW_MMM</td>
<td></td>
<td>RFI</td>
<td>Sherman et al., 2009</td>
</tr>
<tr>
<td>WWMM</td>
<td></td>
<td>RANGLE</td>
<td>Boichard et al., 2003</td>
</tr>
<tr>
<td>WWMM</td>
<td></td>
<td>WWMM</td>
<td>McClure et al., 2010</td>
</tr>
<tr>
<td>WW_rs136155567_27_27.0</td>
<td>LOC104976093, NRG1</td>
<td>BQ</td>
<td>Buitenhuis et al., 2007</td>
</tr>
<tr>
<td>WWMM</td>
<td></td>
<td>SOUND</td>
<td>Buitenhuis et al., 2007</td>
</tr>
<tr>
<td>WWMM</td>
<td></td>
<td>ADFI</td>
<td>Rolf et al., 2012</td>
</tr>
<tr>
<td>WWMM</td>
<td></td>
<td>ADG</td>
<td>Rolf et al., 2012</td>
</tr>
<tr>
<td>WWMM</td>
<td></td>
<td>RFI</td>
<td>Rolf et al., 2012</td>
</tr>
<tr>
<td>WWMM</td>
<td></td>
<td>CALEASE</td>
<td>McClure et al., 2010</td>
</tr>
</tbody>
</table>

1ADG = 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.

2ID = identification.
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

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

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.

<table>
<thead>
<tr>
<th>SNP_BTA</th>
<th>Gene in ±250 kb</th>
<th>Gene ID</th>
<th>Distance, kb</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs43668789_11</td>
<td>GALM</td>
<td>616676</td>
<td>U 217.4</td>
<td>Galactose mutarotase</td>
</tr>
<tr>
<td></td>
<td>SRSF7</td>
<td>507066</td>
<td>U 201.6</td>
<td>Serine and arginine rich splicing factor 7</td>
</tr>
<tr>
<td></td>
<td>GEMIN6</td>
<td>525263</td>
<td>U 160.6</td>
<td>Gem nuclear organelle associated protein 6</td>
</tr>
<tr>
<td></td>
<td>LOC107132913</td>
<td>107132913</td>
<td>U 156.0</td>
<td>Uncharacterized LOC107132913</td>
</tr>
<tr>
<td></td>
<td>DHX57</td>
<td>540993</td>
<td>U 86.1</td>
<td>Dexh-box helicase 57</td>
</tr>
<tr>
<td></td>
<td>MORN2</td>
<td>616607</td>
<td>U 77.8</td>
<td>MORN repeat containing 2</td>
</tr>
<tr>
<td></td>
<td>ARHGEF33</td>
<td>10035703</td>
<td>Cover</td>
<td>Rho guanine nucleotide exchange factor 33</td>
</tr>
<tr>
<td></td>
<td>SOS1</td>
<td>537682</td>
<td>D 17.0</td>
<td>SOS Ras/Rac guanine nucleotide exchange factor 1</td>
</tr>
<tr>
<td></td>
<td>MIR2284Z-2</td>
<td>102465308</td>
<td>D 62.5</td>
<td>Microrna 2284z-2</td>
</tr>
<tr>
<td></td>
<td>LOC104973309</td>
<td>104973309</td>
<td>D 121.0</td>
<td>Ubiquitin-40S ribosomal protein S27a pseudogene</td>
</tr>
<tr>
<td></td>
<td>CDKL4</td>
<td>517478</td>
<td>D 207.4</td>
<td>Cyclin dependent kinase like 4</td>
</tr>
<tr>
<td></td>
<td>LOC782845</td>
<td>782845</td>
<td>D 241.7</td>
<td>60S ribosomal protein L23a pseudogen</td>
</tr>
<tr>
<td>rs136155567_27</td>
<td>LOC104976093</td>
<td>104976093</td>
<td>D 470.9</td>
<td>Uncharacterized LOC104976093</td>
</tr>
<tr>
<td></td>
<td>NRG1</td>
<td>281361</td>
<td>D 567.1</td>
<td>Neuregulin 1</td>
</tr>
</tbody>
</table>

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

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

\(^1\)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 \textit{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. \textit{GALM} is involved in the pathway hexose metabolism, which is part of the carbohydrate metabolism (Thoden \textit{et al.}, 2004). McClure \textit{et al.} (2010) reported a positive association of \textit{GALM} with the weaning weight in Angus cattle. Besides, Shin \textit{et al.} (2014) mentioned that the association between \textit{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 \textit{et al.}, 2010), conformation in dairy cattle breed (Boichard \textit{et al.}, 2003; Buitenhuis \textit{et al.}, 2007), and residual feed intake in Canadian beef synthetic cattle (Sherman \textit{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 \textit{LOC104976093} (uncharacterized LOC104976093) and \textit{NRG1} (neuregulin 1) (Figure 6-Supplements). In this case, \textit{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 \textit{et al.}, 1993; Ieguchi \textit{et al.}, 2010). In cattle, \textit{NRG1} has been highly associated with organs development (Sweeney \textit{et al.}, 2001). Zhao \textit{et al.} (2013) mentioned this gene as an emerging regulator of prolactin secretion, and it could influence weaning weight (Zhao \textit{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


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.