

Genome and chromosome wide association studies for growth traits in Simmental and Simbrah cattle

René Calderón-Chagoya^{1,2}, Vicente Eliezer Vega-Murillo³, Adriana García-Ruiz²,
Ángel Ríos-Utrera⁴, Guillermo Martínez-Velázquez⁵, and Moisés Montaña-Bermúdez^{2,*}

*** Corresponding Author:**

Moisés Montaña-Bermúdez

Tel: +52-55-38-71-8700 Ext. 80220,

E-mail: montano.moises@inifap.gob.mx

¹ Facultad de Medicina Veterinaria y Zootecnia, Universidad Nacional Autónoma de México, Ciudad Universitaria, Ciudad de México 04510, México

² Centro Nacional de Investigación Disciplinaria en Fisiología y Mejoramiento Animal, Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias, Colón, Querétaro 76280, México

³ Facultad de Medicina Veterinaria y Zootecnia, Universidad Veracruzana, Veracruz, Veracruz 91710, México

⁴ Campo Experimental La Posta, Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias, Medellín, Veracruz 94277, México

⁵ Campo Experimental Santiago Ixcuintla, Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias, Santiago Ixcuintla, Nayarit 63570, México

ORCID

René Calderón-Chagoya

<https://orcid.org/0000-0002-4450-8979>

Vicente Eliezer Vega-Murillo

<https://orcid.org/0000-0002-0847-8944>

Adriana García-Ruiz

<https://orcid.org/0000-0002-1074-536X>

Ángel Ríos-Utrera

<https://orcid.org/0000-0003-3108-1133>

Guillermo Martínez-Velázquez

<https://orcid.org/0000-0001-6101-1297>

Moisés Montaña-Bermúdez

<https://orcid.org/0000-0003-3162-6949>

Submitted Nov 25, 2021; Revised Mar 5, 2022;

Accepted Jun 27, 2022

Objective: The objective of this study was to perform genome (genome wide association studies [GWAS]) and chromosome (CWAS) wide association analyses to identify single nucleotide polymorphisms (SNPs) associated with growth traits in registered Simmental and Simbrah cattle.

Methods: The phenotypes were deregressed BLUP EBVs for birth weight, weaning weight direct, weaning weight maternal, and yearling weight. The genotyping was performed with the GGP Bovine 150k chip. After the quality control analysis, 105,129 autosomal SNP from 967 animals (473 Simmental and 494 Simbrah) were used to carry out genotype association tests. The two association analyses were performed per breed and using combined information of the two breeds. The SNP associated with growth traits were mapped to their corresponding genes at 100 kb on either side.

Results: A difference in magnitude of posterior probabilities was found across breeds between genome and chromosome wide association analyses. A total of 110, 143, and 302 SNP were associated with GWAS and CWAS for growth traits in the Simmental-, Simbrah- and joint -data analyses, respectively. It stands out from the enrichment analysis of the pathways for RNA polymerase (POLR2G, POLR3E) and GABAergic synapse (GABRR1, GABRR3) for Simmental cattle and p53 signaling pathway (BID, SERPINB5) for Simbrah cattle.

Conclusion: Only 6,265% of the markers associated with growth traits were found using CWAS and GWAS. The associated markers using the CWAS analysis, which were not associated using the GWAS, represents information that due to the model and priors was not associated with the traits.

Keywords: Genes; Growth Traits; Genome Wide Association Studies (GWAS); Simbrah; Simmental; Single Nucleotide Polymorphism (SNP)

INTRODUCTION

Simmental and Simbrah cattle are some of the most widespread breeds for meat production in Mexico. Growth traits are traditionally included as selection criteria in beef cattle breeding programs, due to their association with meat production and therefore are of great economic importance for both breeders and industry [1]. The most common type of growth trait used in the selection process is the body weight measurement, which can be taken at birth and throughout an animal's life [2]. Growth traits usually present heritability and genetic correlations from moderate to strong [3,4]. Genetic associations are caused by linkage disequilibrium (LD) and pleiotropic effects of genes [5]. The LD is important for experimental designs to increase genome wide association studies (GWAS) efficiency in studied populations. Linkage disequilibrium patterns and scale within and between populations/breeds can be influenced by several factors, such as marker allele

frequencies, selection history, population structure, effective population size, marker type and density, and LD measure used [6]. Additionally, inter-chromosome epistasis effects are expected to be unaffected by LD between the two single nucleotide polymorphisms (SNPs) of each SNP pair but some effects still could have contributions from LD between an SNP adjacent to another SNP that had a significant inter-chromosome epistasis effect [7]. The aim of this research was to implement GWAS and chromosome-wide association (CWAS) analyses to identify SNPs associated with growth traits in registered Simmental and Simbrah cattle.

MATERIALS AND METHODS

The genotype and phenotype of 1,130 animals (547 Simmental and 583 Simbrah), provided by the Mexican Simmental-Simbrah Breeders Association, were used. The phenotypes were deregressed BLUP EBV [8] for birth weight (BW), weaning weight direct (WWD), weaning weight maternal (WWM), and yearling weight (YW).

Blood samples were taken from all the animals in the study. Samples were individually identified and sent to Neogen's GeneSeek Laboratory (Lincoln, NE, USA) for DNA extraction and genotyping using the GGP Bovine 150k chip with 138,962 SNP.

All SNP with a call rate <0.95 and a minor allele frequency <0.05 were excluded. Individuals with a call rate less than 0.95 were also deleted. After the quality control analysis, 105,129 autosomal SNP from 967 animals (473 Simmental and 494 Simbrah) were used to carry out genotype association tests. Additionally, chromosome marker databases were performed to calculate their effect on the different traits individually. Intra-chromosomal LD were evaluated by means of pair-wise coefficients of determination (r^2) in Plink 1.9. Marker pairs were grouped based on their pairwise physical distance into bins of 1 kb, starting from 0 to 5,000 kb. The average r^2 for SNP pairs in each bin was estimated as the arithmetic mean of all r^2 .

A comparison was made between the GWAS and CWAS analyses using a Bayes B model through the BGLR statistical package of the R program. These association analyses were carried out separately for the two breeds and combining the information of both breeds (joint analysis).

For a continuous response (y_i ; $i = 1, \dots, n$) the data equation is represented as $y_i = \eta_i + \varepsilon_i$, where η_i is a linear predictor (the expected value of y_i given predictors) and ε_i are the residuals, independent and normally distributed with mean zero and variance $w_i^2 \sigma_\varepsilon^2$. The linear predictor represents the conditional expectation function, and it is structured as

$$\eta = 1\mu + \sum_{j=1}^J X_j \beta_j,$$

where μ is the intercept, X_j are design matrices for predictors, $X_j = \{x_{ijk}\}$, and β_j are vectors of effects associated with the columns of X_j . For the joint analysis model, X_1 is a design matrix for the effects of breed, β_1 is the corresponding vector of effects, X_2 is the matrix with marker genotypes, and β_2 is the corresponding vector of marker effects. Collecting the above assumptions, we have the following conditional distribution of the data:

$$p(\mathbf{y}|\boldsymbol{\theta}) = \prod_{i=1}^n N \left(y_i | \mu + \sum_{j=1}^J \sum_{k=1}^{K_j} x_{ijk} \beta_{jk}, \sigma_\varepsilon^2 w_i^2 \right)$$

where $\boldsymbol{\theta}$ represents the collection of unknowns, including the intercept, regression coefficients (β_{jk}) and the residual variance.

This analysis uses Markov Chain Monte Carlo (MCMC) methods to calculate posterior mean estimates of marker effects and variances. The chains included 100,000 iterations with the first 25,000 samples used for burn-in.

The models were evaluated with the Gelman–Rubin diagnostic, which evaluates MCMC convergence by analyzing the difference between multiple Markov chains. The convergence is assessed by comparing the estimated between-chains and within-chain variances for each model parameter. For this, the potential scale reduction factor (\hat{R}) was adjusted. The key parameter for this adjustment is the estimated degrees of freedom, d , for a Student-t approximation to the posterior inference based upon the simulations [9],

$$\hat{R}_c = \frac{d+3}{d+1} \hat{R}$$

Large differences between these variances indicate non-convergence ($\hat{R}_c > 1.1$). When the model did not converge, the number of iterations and burn-in were doubled. Once the effects of the markers were estimated, 95% confidence intervals and posterior probabilities for the markers' effects were calculated. Marker effects equal to zero reflect no marker's effect.

The SNP associated with growth traits were mapped at 100 kb pairs on either side, because of the average LD ($r^2 = 0.2$) previously found in beef cattle [10]. With these SNP windows, genes and quantitative trait loci (QTLs) in the same regions were found. After that, pathway enrichment analyses were conducted using the Database for Annotation, Visualization and Integrated Discovery (DAVID). A p-value <0.05 determined by Fisher's exact test was set as the criterion

for significance. The pathway analyses were carried out using the official symbols of genes and the *Bos taurus* species as reference.

RESULTS

The intra-chromosomal LD (r^2) tended to decrease with increasing genomic distance in both breeds and joint data. Simmental showed higher LD than Simbrah in the nearest distances (<100 kb). For larger distances between SNP (>150 kb), Simbrah showed slightly higher average r^2 than Simmental. For adjacent markers (<1 kb), average r^2 were 0.515, 0.587, and 0.469 for joint-, Simmental- and Simbrah-data analyses, respectively (Figure 1).

There were differences also in the average r^2 at a particular marker distance depending on the chromosome. Average r^2 ranged from 0.019 (BTA26) to 0.057 (BTA14) in the joint analysis, from 0.018 (BTA28) to 0.040 (BTA6) in the Simmental-data analysis, and from 0.023 (BTA26) to 0.068 (BTA14) in the Simbrah-data analysis.

Gelman-Rubin's shrink factors of all the models converged to 1 as the number of iterations increased, which indicates that five chains converged to each other. To select those markers associated with a trait, posterior probabilities were used. Even though the order of the most associated markers was similar in both GWAS and CWAS, there was a clear difference between the posterior probabilities obtained with both analyses, e.g., in Simmental, the marker BovineHD1300014404 was associated with YW using both CWAS and GWAS; the difference between posterior probabilities was 0.6805 (Table

1). In almost all cases, when a marker was associated with a trait using GWAS, it was also associated with the same trait using CWAS. In contrast, other markers were associated at chromosome-wide level, but were not associated with a trait at genome-wide level.

In both, GWAS and CWAS, SNP were found to be associated with growth traits in the Simmental population (Supplementary Table S1). A total of 23 (16 for GWAS and 7 for CWAS), 22 (11 for GWAS and 11 for CWAS), 28 (21 for GWAS and 7 for CWAS), and 37 (20 for GWAS and 17 for CWAS) SNP were associated with BW, WWD, WWM, and YW, respectively. Additionally, four of the SNP associated with WWM, and WWD were the same and another four SNP shared the same regions (within 100 kb).

In the Simbrah population, a total of 61 (27 for GWAS and 34 for CWAS), 39 (23 for GWAS and 16 for CWAS), 23 (13 for GWAS and 10 for CWAS), and 20 (14 for GWAS and 6 for CWAS) SNP were associated with BW, WWD, WWM, and YW, respectively (Supplementary Table S1); 9 of these SNPs were the same for WWM and WWD. Another four SNP shared the same regions (within 100 kb) for the same traits.

In the joint analysis, 95 (60 for GWAS and 35 for CWAS), 76 (52 for GWAS and 24 for CWAS), 60 (41 for GWAS and 19 for CWAS), and 71 (42 for GWAS and 29 for CWAS) SNP were associated with BW, WWD, WWM and YW, respectively (Supplementary Table S1). Also, 19 and 1 of the SNPs associated with WWM and YW were the same for WWD. Another 3, 1, and 1 SNPs for WWM, BW, and YW shared the same regions (within 100 kb) with SNPs associated with WWD. Within the SNP windows, QTLs were previously associated with other traits, 183, 162, 18, and 302 within the regions associated with BW, WWD, WWM and YW, respectively (Supplementary Table S2). In the same way, a total of 51, 61, 25, and 40 genes were found within these regions (Supplementary Table S2).

In each type of analysis, markers associated with growth traits were found in at least 10 chromosomes, however, there is no clear difference between which chromosome had more associated regions. Displaying all SNP for Simmental, Simbrah and joint model, the SNP for WWD matched 34 and 2 SNP for BW, WWM, and YW, respectively. Also, 3, 2, and 1 SNP windows for BW matched with the SNPs for WWD, WWM, and YW; 5 and 1 SNP windows for WWD matched with the SNPs for WWM and YW; finally, 1 SNP window matched between WWM and YW. Additionally, 14 and 1 of the SNPs associated with WWM and YW were the same for WWD. However, these SNP windows are not close enough to match three or more traits at the same time.

Through the database QTL_ARS_UCD1, 280, 271, 156, and 402 QTLs previously described were found within the SNP windows for BW, WWD, WWM, and YW, respectively

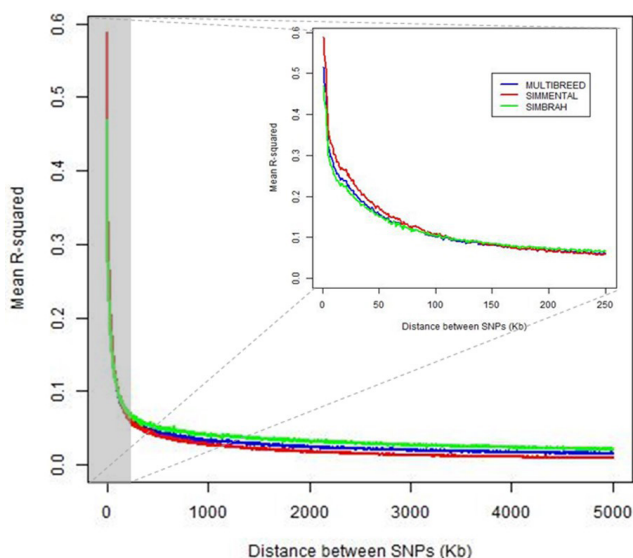


Figure 1. Average coefficients of determination (r^2) for distance bins of 1 kb for Simmental-, Simbrah- and joint- data analyses plotted against physical distance (kb) between single nucleotide polymorphism.

Table 1. Single nucleotide polymorphisms (SNP) with the highest posterior probabilities (>0.15) associated with growth traits in both genome-wide (GWAS) and chromosome-wide analysis (CWAS)

| SNP | BW | | WWD | | WWM | | YW | |
|--------------------|-------|-------|-------|-------|-------|-------|-------|-------|
| | GWAS | CWAS | GWAS | CWAS | GWAS | CWAS | GWAS | CWAS |
| Joint analysis | | | | | | | | |
| BovineHD0400015064 | | | | | | | 0.155 | 0.660 |
| BovineHD0700022164 | 0.269 | 0.552 | | | | | | |
| BovineHD1100018479 | | | 0.448 | 0.794 | 0.405 | 0.907 | | |
| BovineHD1200007354 | | | 0.188 | 0.226 | | | | |
| BovineHD1900010762 | | | 0.428 | 0.964 | | | | |
| BovineHD2000005310 | | | | | | | 0.287 | 0.914 |
| BovineHD2200002151 | 0.750 | 0.860 | | | | | | |
| BovineHD2900012617 | 0.664 | 0.338 | | | | | | |
| BTB-00218987 | 0.251 | 0.482 | | | | | | |
| BTB-00397480 | | | 0.225 | 0.476 | | | | |
| Simbrah analysis | | | | | | | | |
| ARS-BFGL-NGS-3343 | 0.288 | 0.345 | | | | | | |
| ARS-BFGL-NGS-40907 | | | 0.253 | 0.318 | 0.152 | | | |
| ARS-BFGL-NGS-75279 | | | 0.210 | 0.476 | | 0.157 | | |
| BovineHD0100005053 | 0.259 | 0.750 | | | | | | |
| BovineHD0200037022 | | | | 0.272 | 0.150 | | | |
| BovineHD0300022871 | | | | 0.368 | 0.200 | 0.687 | | |
| BovineHD0400014726 | | | 0.171 | | 0.291 | 0.236 | | |
| BovineHD0700030669 | | | | | | | 0.170 | 0.694 |
| BovineHD1100018735 | | | | | | | 0.241 | 0.903 |
| BovineHD2000005310 | | | | | | | 0.220 | 0.503 |
| Simmental analysis | | | | | | | | |
| BovineHD1100018476 | | | 0.171 | 0.675 | | | | |
| BovineHD1100028536 | | | | | | | 0.338 | 0.821 |
| BovineHD1300014404 | | | | | | | 0.151 | 0.831 |
| BovineHD2100010794 | | | 0.153 | 0.590 | | | | |
| BovineHD2100012099 | 0.543 | 0.609 | | | | | | |
| BovineHD2800009062 | | | | | | | 0.179 | 0.471 |

BW, birth weight; WWD, weaning weight direct; WWM, weaning weight maternal; YW, yearling weight.

(Supplementary Table S2). For all the traits, QTLs associated with conformation, health, meat, carcass, milk, and reproduction traits in cattle were found. Over the database ARS-UCD1.2, 131, 102, 64, and 98 genes were found within the SNP windows for BW, WWD, WWM, and YW, respectively (Supplementary Table S2). Four of these genes were found within SNP windows for WWD and YW.

The posterior probability for gene inclusion is always greater than or equal to the probability that any SNP is included [11]. For this reason, genes within SNP windows were used to search for networks that could be associated with growth traits. Functional enrichment analysis was carried out to identify genes that are over-represented in a large group of genes and may have a connection with the studied phenotypes (Table 2).

DISCUSSION

The LD estimates at various distances were of the magnitude

of those reported by Villa-Angulo et al [12] in several dairy and beef cattle breeds using less dense SNP panels. The difference in the decline of the average r^2 between Simmental and Simbrah could be an effect of the indicine breeds. It has been observed that indicine breeds had lower r^2 at short distances and higher r^2 at longer distances between markers than taurine breeds [13]. Higher LD in taurine breeds has been attributed to smaller effective population size and stronger genetic bottleneck during breed formation [14].

Means of r^2 were obtained for each chromosome averaged across breeds ranging from 0.019 (BTA26) to 0.057 (BTA14) in the joint model, from 0.018 (BTA28) to 0.040 (BTA6) in Simmental, and from 0.023 (BTA26) to 0.068 (BTA14) in Simbrah. In the present study, the higher LD values detected in some chromosomes in comparison to others can be indicative of the presence of QTLs affecting traits that have been under intense selection in both breeds [6]. A wide variation in autosomal recombination rates can lead to a marked diversity in the pattern of LD in different genomic regions and

Table 2. Reactome and KEGG pathways significantly enriched using genes associated with growth traits

| Data | Trait | Category | Pathway | Count | p-value | Candidate genes |
|-----------|-------|----------|---|--|---------|-----------------------|
| Joint | WWD | Reactome | Hemostasis-Thrombin signalling through proteinase activated receptors | 2 | 0.02 | <i>GNA11, GNA15</i> |
| | | | Hemostasis-Thromboxane signalling through TP receptor | 2 | 0.0096 | <i>GNA11, GNA15</i> |
| | | | Hemostasis-ADP signalling through P2Y purinoceptor 1 | 2 | 0.012 | <i>GNA11, GNA15</i> |
| | | | Hemostasis-Platelet homeostasis-Platelet sensitization by LDL | 2 | 0.011 | <i>GNA11, GNA15</i> |
| | | | Metabolism-Fatty Acids bound to GPR40 regulate insulin secretion | 2 | 0.0096 | <i>GNA11, GNA15</i> |
| Simmental | BW | KEGG | RNA polymerase | 2 | 0.035 | <i>POLR2G, POLR3E</i> |
| | | | Reactome | Metabolism of proteins-N-glycan trimming in the endoplasmic reticulum and Calnexin/ Calreticulin cycle | 2 | 0.017 |
| | WWD | KEGG | Nicotine addiction | 2 | 0.022 | <i>GABRR1, GABRR2</i> |
| | | | GABAergic synapse | 2 | 0.045 | <i>GABRR1, GABRR2</i> |
| | | | Morphine addiction | 2 | 0.048 | <i>GABRR1, GABRR2</i> |
| Simbrah | BW | KEGG | p53 signaling pathway | 2 | 0.046 | <i>BID, SERPINB5</i> |

KEGG, Kyoto encyclopedia of genes and genomes; *GNA11*, G protein subunit alpha 11; *GNA15*, G protein subunit alpha 15; WWD, weaning weight direct; BW, birth weight; *POLR2G*, RNA polymerase II subunit G; *POLR3E*, RNA polymerase III subunit E; *UBXN1*, UBX domain protein 1; *GANAB*, glucosidase II alpha subunit; *GABRR1*, gamma-aminobutyric acid type A receptor subunit rho1; *GABRR2*, gamma-aminobutyric acid type A receptor subunit rho2; *BID*, BH3 interacting domain death agonist; *SERPINB5*, serpin family B member 5.

chromosomes [15]. The causes may have acted differently at specific genomic regions at singular locations among the Simmental and Simbrah populations. The different LD patterns across chromosomes have been assumed that exist, and therefore, it is also expected different genomic inflation factors across chromosomes [16].

The posterior probabilities do not define with clarity the extent of association of an SNP with a trait, however, if we compare with frequentist models, studies have been carried out where it is shown that around half of the published associations with $p < 5 \times 10^{-7}$ had posterior probabilities less than 0.5 [17]. The posterior probabilities are being conditioned on the model and the priors, however, while posterior probabilities provide a measure of evidence for hypotheses for the marker effects, it is difficult to judge them separately, as individual model probabilities may be “diluted” as the number of markers grows receiving small probability (both prior and posterior) [11,18]; this could have been what affected the posterior probabilities of the GWAS and CWAS; if this is the case, a lot of useful information could be missing.

The change in posterior probabilities was probably due to the density of the markers used in each of the three analyses, with the GWAS giving lower values compared to those of the CWAS [19]. However, a study applying a mixed linear association model with a leave-one-chromosome-out approach, suggests that even if the genomic inflation factors do not differ a lot between the different SNP densities, genomic inflation factors varied largely across the chromosomes [16]. An explanation might be that there was a different level of association between the SNP on the chromosome and the trait of interest, also, because using the total number of SNPs can result in too conservative thresholds since it violates the

assumption of independence between tests [20,21].

For all growth traits, markers were found in regions previously associated with QTLs for production, reproduction, health, and conformation traits (Supplementary Table S2). However, it is of greater interest to focus on the QTLs that are correlated with growth traits.

In the Simmental-data analysis, within the SNP windows, QTL were previously associated with other traits; in total, there were 69, 30, 135, and 71 QTLs within the regions associated with BW, WWD, WWM, and YW, respectively (Supplementary Table S2). A total of 60, 17, 34, and 50 genes were associated with BW, WWD, WWM, and YW, respectively (Supplementary Table S2). For BW, in the associated SNP windows, some QTLs were previously reported with the length of productive life, body depth, net merit, muscle phosphorus and potassium content, shear force, and tenderness score in Holstein, Angus, and Nelore. For WWD, one SNP window on BTA 10 were reported QTLs associated with body weight in Charolais and Gelbvieh cattle. In the case of WWM, inside the SNP windows QTLs were previously associated with average daily gain, body weight gain, body depth, rump width, body weight, carcass weight, fat thickness, hip height, longissimus muscle area, marbling score, metabolic body weight, residual feed intake, withers height, and dry matter intake. For YW, it was found in SNP windows QTLs previously associated with length of productive life, net merit, and dry matter intake in Holstein cattle.

In Simbrah cattle, inside the SNP windows, 28, 79, 3, and 29 QTL were previously associated with other traits within the regions associated with BW, WWD, WWM, and YW, respectively (Supplementary Table S2). Also, a total of 20, 24,

5, and 8 genes were found within these regions (Supplementary Table S2). Within regions associated with BW QTLs were previously found for maturity rate in Brahman cattle and net merit and length of productive life in Holstein cattle. For WWD, in the SNP windows QTLs were previously associated with body weight (birth) in Brangus cattle, weaning weight in Blanco Orejinegro cattle and lean meat yield in Holstein cattle. For WWM, on BTA 24, QTL previously associated with bodyweight and maturity rate were found. For YW, in a SNP window of BTA 14 a QTL previously associated with birth weight in Charolais and Chianina cattle and carcass weight in Hanwoo cattle was found.

From the results of the joint model, several coincidences were found within the regions associated with BW, WWD, and YW, with the databases of QTLs. For BW, in the associated SNP windows, some QTLs previously reported with maturity rate in Brahman cattle, body weight (birth) in Charolais and Chianina cattle, yearling weight in Charolais and Gelbvieh cattle, subcutaneous fat in Hanwoo cattle, body depth, lean meat yield, length of productive life, net merit, PTA type and rump width in Holstein cattle, and dry matter intake and metabolic weight in beef cattle. In the case of WWD inside the SNP windows, QTLs were previously associated with average daily feed intake, carcass weight, fat thickness, marbling score, residual feed intake, and, average daily gain in beef cattle, body weight (yearling) in Charolais, and Chianina cattle, maturity rate in Brahman cattle, tenderness score in Angus cattle and, lean meat yield, growth index, average daily gain, length of productive life and net merit in Holstein cattle. For WWM, inside the SNP windows were reported QTLs associated with muscle sodium content in Angus cattle, maturity rate in Brahman cattle, and dry matter intake in Holstein cattle. For YW it was found in SNP windows QTLs previously associated with marbling score, muscle creatine content, tenderness score in Angus cattle, bodyweight in Charolais, Chianina, and Gelbvieh cattle, carcass weight in Wagyu cattle, dry matter intake, dressed carcass, and body mass in a beef cattle population, and body depth, lean meat yield, length of productive life, net merit and rump width and lactation persistence in Holstein cattle.

The joint analysis of two breeds increases the statistical power to detect more significant SNPs rather than a single analysis. Also, the Simmental and Simbrah analyses give different significant genes. These differences can be associated with the genomic architecture changes with hybridization and subsequent inter-se mating during the formation of a composite breed, this means that alleles at some loci increase in frequency more than others in the newly hybridized population [22]. Additionally, differences were found between the enrichment analyzes, mainly because no similar genes were found between the different association analyses. Furthermore, the more diverse the number of genes, the greater

the number of pathways that were found.

In the joint analysis, inside the SNP windows, there are some genes relevant to growth traits (Supplementary Table S2). Among these stand out special AT-rich sequence-binding protein-1 (*SATB1*), prominin mouse-like 1 (*PROM1*), CUB and Sushi multiple domains 1 (*CSMD1*), phosphatidyserine synthase 1 (*PTDSS1*), and ubiquinol-cytochrome c reductase binding protein (*UQCRB*) (candidate genes for BW).

SATB1 has been associated with the concentration of triiodothyronine, a hormone linked with physiological processes, including growth and development [23]. Also, *SATB1* has shown differences in its expression in fetal adipose tissues, depending on the maternal diet [24]. *PROM1* has been selected as a candidate gene for BW in goats, this is a protein-coding gene, which plays a role in cell differentiation, proliferation, and apoptosis [25]. In Hanwoo cattle, *CSMD1* was more highly expressed in muscle samples from animals with increasing carcass weight in intramuscular fat and eye muscle area [26]. Additionally, *PTDSS1* and *UQCRB* tend to be expressed more highly in muscle with increasing intramuscular fat content [26].

Reactome pathways for WWD in the joint analysis included the metabolism through the regulation of insulin secretion by fatty acids bound to G protein-coupled receptor 40 (*GPR40*) fatty acids augment the glucose-triggered secretion of insulin through two mechanisms: intracellular metabolism and activation of free fatty acid receptor 1 (*FFAR1*).

Also, in the Simmental-data analysis, for BW, a reactome pathway associated with the metabolism of proteins-N-glycan trimming in the endoplasmic reticulum and Calnexin/Calreticulin cycle was identified; in this process, the N-glycan is progressively trimmed off by the three glucoses and some of the mannoses before the protein is transported to the cis-Golgi.

In Simmental cattle, inside the SNP windows, there are some genes relevant to growth traits (Supplementary Table S2). Among these stand out *SATB1* (candidate gene for WWD), proenkephalin (*PENK*), collagen type IV alpha 1 chain (*COL4A1*), short chain dehydrogenase/reductase family 16C member 6 (*SDR16C6*), and cell division cycle 5 like (*CDC5L*) (candidate genes for WWM).

PENK has been associated as a candidate gene for growth traits along with RB1 inducible coiled-coil 1 (*RBICC1*), neuropeptides B and W receptor 1 (*NPBWR1*), and pleiomorphic adenoma gene 1 (*PLAG1*) in Nellore cattle, according to these studies, *PLAG1* seems to be the major gene due to its role in regulating insulin-like growth factors, and *RBICC1*, *NPBWR1*, and *PENK* are also involved in processes that can contribute to determining the uniformity of growth traits, like YW [27]. *COL4A1* is related to developmental biology, protein digestion, and protein absorption have been indicated

as a candidate gene for weaning weight and YW [28]. *SDR16C6* has been associated as a candidate gene for weaning weight and daily weight gain between birth and weaning in Blanco Orejinerio cattle [29]. *CDC5L* has been associated as a candidate gene of muscular growth and homeostasis during puberty in conjunction with *MYC* proto-oncogene (*MYC*), transcription factor 3 (*TCF3*), *RUNX* family transcription factor 2 (*RUNX2*), activating transcription factor 2 (*ATF2*), and cAMP responsive element binding protein 1 (*CREB1*) [30].

The RNA polymerase pathway was associated with Simmental cattle for BW and take part in the transcription in the genetic information processing. In a study the RNA polymerase pathway was observed downregulated in overfed moderate-energy diet (OVE) cows; also, the gene *POLR2G* (polymerase II gene) and other polymerases III genes were affected. To lower polymerase II gene expression, OVE cows also experienced suppression of the RNA transport pathway. RNA transport allows mRNA transcribed in the nucleus to be processed and translated later in the cytoplasm [31]. This result does not necessarily mean that there was less overall transcription. In eukaryotes, there are 3 distinct RNA polymerases, these transcription complexes are composed of heterogeneous subunits, which can individually affect the transcription complex [32].

The GABAergic synapse pathway was associated with WWD in Simmental cattle. GABA is a neurotransmitter widely distributed in the central nervous system, which is synthesized from glutamate through decarboxylation [39] and plays an important role in regulating feeding behavior in the hypothalamus. Other studies have found that this pathway was significantly associated with live weight in Simmental cattle [33]. Additionally, it was suggested that neuronal sensitivity to GABA is related to the control of feeding behavior in ruminant animals [34].

An interesting term that has been found in Simmental cattle for WWD is nicotine and morphine addiction. Both terms are found as significant Kyoto encyclopedia of genes and genomes pathways associated with substance dependence in humans extrapolated to bovines. Other studies have observed these pathways in their analysis [33,35,36]. However, given the origin, all agree that the pathway regulation is ambiguous and requires further validation.

In Simbrah cattle, inside the SNP windows, there are some genes relevant to growth traits (Supplementary Table S2). Among these stand out vacuolar protein sorting 4 homolog B (*VPS4B*) (candidate gene for BW), cadherin 20 (*CDH20*) (candidate gene for WWM), hedgehog acyltransferase (*HHAT*), phosphodiesterase 4B (*PDE4B*), tripartite motif containing 63 (*TRIM63*), high mobility group AT-hook 2 (*HMGA2*) (candidate genes for WWD), and thymocyte selection associated high mobility group box (*TOX*) (candidate gene for YW).

In pigs, *VPS4B* and *CDH20* have been described as a candidate gene that composes the underlying genetic architecture of porcine growth and fatness traits, this gene is crucial for the degradation of membrane receptors, regulation of epidermal growth factor receptors, and insulin receptors [37].

HHAT is a gene previously associated with weaning weight in Zebú cattle [38]. *PDE4B* encodes the phosphodiesterase enzyme type 4 that hydrolyses the cyclic adenosine monophosphate, which is related to energy modulation processes in the body, and is linked with lipolysis control, regulating body composition, also this gene has been associated with average daily gain [39]. *TRIM63* is part of the ubiquitin-proteasome system in the main proteolytic pathway in muscle, and the muscle-specific ligases tripartite motif-containing, also there is a supposition that *TRIM63* (MuRF-1) may play a role in the control of protein degradation and probably also contributes to skeletal muscle metabolism [40]. *HMGA2* has been detected for BW in Brangus, the *HMGA* proteins are architectural transcription factors that regulate the transcription of a variety of genes and direct cellular growth, proliferation, and differentiation [41], also the regulation of insulin like growth factor 2 (*IGF2*) by *HMGA2* has been proved to occur directly or through increased expression of *PLAG1* [42].

The *TOX* gene has previously been linked to *PLAG1*, coiled-coil-helix-coiled-coil-helix domain containing 7 (*CHCHD7*), short chain dehydrogenase/reductase family 16C member 5 (*SDR16C5*), *SDR16C6*, *PENK*, family with sequence similarity 110 member B (*FAM110B*), cytochrome P450 family 7 subfamily A member 1 (*CYP7A1*), and syndecan binding protein (*SDCBP*) as candidate genes for carcass weight in Hanwoo cattle. According to them, a denser LD structure was found in and around *TOX* gene rather than a region that surrounds *PLAG1* gene, this result might be due to a multigene effect in which multiple genes in the same QTL region are affecting correlated traits in cattle [43].

In Simbrah, the p53 signaling pathway was significantly associated with BW. This pathway is induced by several stress signals (DNA damage, oxidative stress, and activated oncogenes). The p53 protein is employed as a transcriptional activator of p53-regulated genes. Also, p53 is a gene that has been proposed as a part of a network that influences puberty, making supports the relevance of tumor-related genes for puberty [44]. There is evidence of the correlation of pubertal traits (standardized age at first oestrus and scrotal circumference) with growth characteristics, such as YW and the maternal component of weaning weight and BW [45].

CONCLUSION

Only 6,265% of the markers associated with growth traits were found using CWAS and GWAS. The associated markers

found just in the CWAS could be used for the identification of candidate genes. No significantly associated regions were found between breeds. Although Simbrah is a synthetic breed derived from Simmental, no common regions were found, however, in the joint analysis, some common regions were found.

These regions may be useful in providing insight into growth traits in Simmental and Simbrah cattle with related phenotypic measurements. Also, candidate genes helped identify gene pathways through enrichment analysis. These pathways can help us understand how they are connected to growth traits.

CONFLICT OF INTEREST

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

FUNDING

This article was funded by the National Institute for Forestry, Agricultural and Livestock Research (INIFAP) of Mexico, through the project: 854434754. Also, the provision of information by the Mexican Simmental-Simbrah Breeders Association is greatly appreciated.

SUPPLEMENTARY MATERIAL

Supplementary file is available from: <https://doi.org/10.5713/ab.21.0517>

Supplementary Table S1. Marker effects (β) and their posterior probabilities (PP) of SNPs associated with growth traits with the use of genome-wide association (GWA) as well as a chromosome-wide association (CWA) in simmental and simbrah cattle

Supplementary Table S2. QTLs and genes previously described within the regions associated with growth traits in simmental and simbrah cattle

REFERENCES

1. Barwick SA, Henzell AL. Development successes and issues for the future in deriving and applying selection indexes for beef breeding. *Aust J Exp Agric* 2005;45:923-33. <https://doi.org/10.1071/EA05068>
2. Buzanskas ME, Grossi DA, Ventura RV, et al. Genome-wide association for growth traits in Canchim beef cattle. *PLoS One* 2014;9:e94802. <https://doi.org/10.1371/journal.pone.0094802>
3. Buzanskas ME, Grossi DA, Baldi F, et al. Genetic associations between stayability and reproductive and growth traits in Canchim beef cattle. *Livest Sci* 2010;132:107-12. <https://doi.org/10.1016/j.livsci.2010.05.008>
4. Baldi F, Albuquerque LG, Alencar MM. Random regression models on Legendre polynomials to estimate genetic parameters for weights from birth to adult age in Canchim cattle. *J Anim Breed Genet* 2010;127:289-99. <https://doi.org/10.1111/j.1439-0388.2010.00853.x>
5. Terakado APN, Costa RB, de Camargo GME, et al. Genome-wide association study for growth traits in Nelore cattle. *Animal* 2018;12:1358-62. <https://doi.org/10.1017/S1751731117003068>
6. Biegelmeyer P, Gúlias-Gomes CC, Caetano AR, Steibel JP, Cardoso FF. Linkage disequilibrium, persistence of phase and effective population size estimates in Hereford and Braford cattle. *BMC Genet* 2016;17:32. <https://doi.org/10.1186/s12863-016-0339-8>
7. Prakapenka D, Liang Z, Jiang J, Ma L, Da Y. A large-scale genome-wide association study of epistasis effects of production traits and daughter pregnancy rate in u.S. Holstein cattle. *Genes (Basel)* 2021;12:1089. <https://doi.org/10.3390/genes12071089>
8. Garrick DJ, Taylor JF, Fernando RL. Deregressing estimated breeding values and weighting information for genomic regression analyses. *Genet Sel Evol* 2009;41:55. <https://doi.org/10.1186/1297-9686-41-55>
9. Brooks SP, Gelman A. General methods for monitoring convergence of iterative simulations. *J Comput Graph Stat* 1998;7:434-55. <https://doi.org/10.1080/10618600.1998.10474787>
10. Akanno EC, Chen L, Abo-Ismael MK, et al. Genome-wide association scan for heterotic quantitative trait loci in multi-breed and crossbred beef cattle. *Genet Sel Evol* 2018;50:48. <https://doi.org/10.1186/s12711-018-0405-y>
11. Wilson MA, Iversen ES, Clyde MA, Schmidler SC, Schildkraut JM. Bayesian model search and multilevel inference for SNP association studies. *Ann Appl Stat* 2010;4:1342-64. <https://doi.org/10.1214/09-AOAS322>
12. Villa-Angulo R, Matukumalli LK, Gill CA, Choi J, Van Tassell CP, Grefenstette JJ. High-resolution haplotype block structure in the cattle genome. *BMC Genet* 2009;10:19. <https://doi.org/10.1186/1471-2156-10-19>
13. Gibbs RA, Taylor JF, Van Tassell CP, et al. Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. *Science* 2009;324:528-32. <https://doi.org/10.1126/science.1167936>
14. Porto-Neto LR, Kijas JW, Reverter A. The extent of linkage disequilibrium in beef cattle breeds using high-density SNP genotypes. *Genet Sel Evol* 2014;46:22. <https://doi.org/10.1186/1297-9686-46-22>
15. Cañas-Álvarez JJ, Mouresan EF, Varona L, et al. Linkage disequilibrium, persistence of phase, and effective population size in Spanish local beef cattle breeds assessed through a high-density single nucleotide polymorphism chip. *J Anim Sci* 2016;94:2779-88. <https://doi.org/10.2527/jas.2016-0425>

16. den Berg S, Vandenplas J, Eeuwijk FA, Lopes MS, Veerkamp RF. Significance testing and genomic inflation factor using high-density genotypes or whole-genome sequence data. *J Anim Breed Genet* 2019;136:418-29. <https://doi.org/10.1111/jbg.12419>
17. Ball RD. Designing a GWAS: power, sample size, and data structure. In: *Methods in molecular biology*. Totowa, NJ, USA; Humana Press Inc.; 2013. Vol. 1019, pp. 37-98.
18. Clyde M, George EI. Model uncertainty. *Statist Sci* 2004;19: 81-94. <https://doi.org/10.1214/088342304000000035>
19. Cheung SH, Beck JL. Calculation of posterior probabilities for bayesian model class assessment and averaging from posterior samples based on dynamic system data. *Comput Civ Infrastruct Eng* 2010;25:304-21. <https://doi.org/10.1111/j.1467-8667.2009.00642.x>
20. Duggal P, Gillanders EM, Holmes TN, Bailey-Wilson JE. Establishing an adjusted p-value threshold to control the family-wide type 1 error in genome wide association studies. *BMC Genomics* 2008;9:516. <https://doi.org/10.1186/1471-2164-9-516>
21. Nicodemus KK, Liu W, Chase GA, Tsai YY, Fallin MD. Comparison of type I error for multiple test corrections in large single-nucleotide polymorphism studies using principal components versus haplotype blocking algorithms. *BMC Genet* 2005;6:S78. <https://doi.org/10.1186/1471-2156-6-S1-S78>
22. Paim TP, Hay EHA, Wilson C, et al. Dynamics of genomic architecture during composite breed development in cattle. *Anim Genet* 2020;51:224-34. <https://doi.org/10.1111/age.12907>
23. Tang J, Hu W, Chen S, et al. The genetic mechanism of high prolificacy in small tail han sheep by comparative proteomics of ovaries in the follicular and luteal stages. *J Proteomics* 2019;204:103394. <https://doi.org/10.1016/j.jprot.2019.103394>
24. Peñagaricano F, Wang X, Rosa GJM, Radunz AE, Khatib H. Maternal nutrition induces gene expression changes in fetal muscle and adipose tissues in sheep. *BMC Genomics* 2014; 15:1034. <https://doi.org/10.1186/1471-2164-15-1034>
25. Zhang L, Wang F, Gao G, et al. Genome-wide association study of body weight traits in Inner Mongolia cashmere goats. *Front Vet Sci* 2021;8:752746. <https://doi.org/10.3389/fvets.2021.752746>
26. Lee SH, van der Werf JHJ, Kim NK, et al. QTL and gene expression analyses identify genes affecting carcass weight and marbling on BTA14 in Hanwoo (Korean Cattle). *Mamm Genome* 2011;22:589. <https://doi.org/10.1007/s00335-011-9331-9>
27. Iung LHS, Mulder HA, Neves HHR, Carvalheiro R. Genomic regions underlying uniformity of yearling weight in Nellore cattle evaluated under different response variables. *BMC Genomics* 2018;19:619. <https://doi.org/10.1186/s12864-018-5003-4>
28. Weng Z, Su H, Saatchi M, et al. Genome-wide association study of growth and body composition traits in Brangus beef cattle. *Livest Sci* 2016;183:4-11. <https://doi.org/10.1016/j.livsci.2015.11.011>
29. Londoño-Gil M, Rincón Flórez JC, Lopez-Herrera A, Gonzalez-Herrera LG. Genome-wide association study for growth traits in blanco orejiner (bon) cattle from Colombia. *Livest Sci* 2021;243:104366. <https://doi.org/10.1016/j.livsci.2020.104366>
30. Lau LY, Nguyen LT, Reverter A, et al. Gene regulation could be attributed to TCF3 and other key transcription factors in the muscle of pubertal heifers. *Vet Med Sci* 2020;6:695-710. <https://doi.org/10.1002/VMS3.278>
31. Kireeva ML, Kashlev M, Burton ZF. RNA polymerase structure, function, regulation, dynamics, fidelity, and roles in gene expression. *Chem Rev* 2013;113:8325-30. <https://doi.org/10.1021/cr400436m>
32. Agrawal A, Khan MJ, Graugnard DE, et al. Prepartal energy intake alters blood polymorphonuclear leukocyte transcriptome during the peripartal period in Holstein cows. *Bioinform Biol Insights* 2017;11:1-17. <https://doi.org/10.1177/1177932217704667>
33. Fan H, Wu Y, Zhou X, et al. Pathway-based genome-wide association studies for two meat production traits in Simmental cattle. *Sci Rep* 2016;5:18389. <https://doi.org/10.1038/srep18389>
34. Seoane JR, Dumont F, Girard CL, Bédard L, Matte JJ. Effects of intraventricular injections of gamma-aminobutyric acid and related substances on feeding behavior in satiated sheep. *Can J Physiol Pharmacol* 1984;62:1296-9. <https://doi.org/10.1139/y84-217>
35. Yoon J, Kim H. Multi-tissue observation of the long non-coding RNA effects on sexually biased gene expression in cattle. *Asian-Australas J Anim Sci* 2019;32:1044-51. <https://doi.org/10.5713/ajas.18.0516>
36. Sun Y, Wang C, Sun X, Guo, M. Comparative proteomics of whey and milk fat globule membrane proteins of Guanzhong goat and Holstein cow mature milk. *J Food Sci* 2019;84:244-53. <https://doi.org/10.1111/1750-3841.14428>
37. Zhang Z, Chen Z, Diao S, et al. Identifying the complex genetic architecture of growth and fatness traits in a Duroc pig population. *J Integr Agric* 2021;20:1607-14. [https://doi.org/10.1016/S2095-3119\(20\)63264-6](https://doi.org/10.1016/S2095-3119(20)63264-6)
38. Martínez R, Gómez Y, Rocha, J. Genome-wide association study on growth traits in Colombian creole breeds and crossbreeds with Zebu cattle. *Genet Mol Res* 2014;13:6420-32. <https://doi.org/10.4238/2014.August.25.5>
39. Santana MHA, Utsunomiya YT, Neves HHR, et al. Genome-wide association study for feedlot average daily gain in Nellore cattle (*Bos indicus*). *J Anim Breed Genet* 2014;131:210-6. <https://doi.org/10.1111/JBG.12084>
40. Yang Y, Sadri H, Prehn C, et al. Proteasome activity and expression of mammalian target of rapamycin signaling

- factors in skeletal muscle of dairy cows supplemented with conjugated linoleic acids during early lactation. *J Dairy Sci* 2020;103:2829-46. <https://doi.org/10.3168/jds.2019-17244>
41. Saatchi M, Schnabel RD, Taylor JF, Garrick DJ. Large-effect pleiotropic or closely linked QTL segregate within and across ten US cattle breeds. *BMC Genomics* 2014;15:442. <https://doi.org/10.1186/1471-2164-15-442>
 42. Abi Habib W, Brioude F, Edouard T, et al. Genetic disruption of the oncogenic HMGA2–PLAG1–IGF2 pathway causes fetal growth restriction. *Genet Med* 2018;20:250-8. <https://doi.org/10.1038/gim.2017.105>
 43. Lee SH, Choi BH, Lim D, et al. Genome-Wide Association Study Identifies Major Loci for Carcass Weight on BTA14 in Hanwoo (Korean Cattle). *PLoS One* 2013;8:e74677. <https://doi.org/10.1371/journal.pone.0074677>
 44. Fortes MRS, Reverter A, Zhang Y, et al. Association weight matrix for the genetic dissection of puberty in beef cattle. *Proc Natl Acad Sci* 2010;107:13642-7. <https://doi.org/10.1073/pnas.1002044107>
 45. Morris CA, Wilson JA, Bennett GL, Cullen NG, Hickey SM, Hunter JC. Genetic parameters for growth, puberty, and beef cow reproductive traits in a puberty selection experiment. *NZ J Agric Res* 2000;43:83-91. <https://doi.org/10.1080/00288233.2000.9513411>