Identification of Candidate Variants Associated With Bone Weight Using Whole Genome Sequence in Beef Cattle

Qunhao Niu[†], Tianliu Zhang[†], Ling Xu, Tianzhen Wang, Zezhao Wang, Bo Zhu, Xue Gao, Yan Chen, Lupei Zhang, Huijiang Gao, Junya Li* and Lingyang Xu*

Key Laboratory of Animal Genetics Breeding and Reproduction, Ministry of Agriculture and Rural Affairs, Institute of Animal Sciences, Chinese Academy of Agricultural Sciences, Beijing, China

Bone weight is critical to affect body conformation and stature in cattle. In this study, we conducted a genome-wide association study for bone weight in Chinese Simmental beef cattle based on the imputed sequence variants. We identified 364 variants associated with bone weight, while 350 of them were not included in the Illumina BovineHD SNP array, and several candidate genes and GO terms were captured to be associated with bone weight. Remarkably, we identified four potential variants in a candidate region on BTA6 using Bayesian fine-mapping. Several important candidate genes were captured, including *LAP3*, *MED28*, *NCAPG*, *LCORL*, *SLIT2*, and *IBSP*, which have been previously reported to be associated with carcass traits, body measurements, and growth traits. Notably, we found that the transcription factors related to *MED28* and *LCORL* showed high conservation across multiple species. Our findings provide some valuable information for understanding the genetic basis of body stature in beef cattle.

Keywords: genetic architecture, imputed sequence variants, bone weight, Bayesian fine-mapping, beef cattle

INTRODUCTION

Cattle, as one of the most important farm animals, provide numerous meat products for high-quality protein. Carcass merit traits are a key factor that directly affects meat yield, and their performance is highly connected with stature size (Sieber et al., 1988; Albertí et al., 2008; Ozkaya and Bozkurt 2008). Bone weight can reflect the size of stature and the skeleton frame (Berg and Butterfield 1966; Chumlea et al., 2002; Conroy et al., 2010) and is involved in the respiratory disease and feed efficiency in cattle (Snowder et al., 2007; Mader et al., 2009). Meanwhile, many economically important traits (carcass weight, meat yield, etc.) were significantly associated with bone weight (Pabiou et al., 2012). Therefore, elucidation of the genetic basis for bone weight can provide valuable information into understanding the bone development (Browning and Browning 2009), as well as exploring the potential the molecular mechanism of stature size in cattle.

Over the past few years, genome-wide association studies (GWAS) have been widely utilized to identify quantitative trait loci (QTLs) associated with complex traits in cattle (Wu et al., 2013; Bouwman et al., 2018; Fang and Pausch 2019). Several QTLs relevant to stature, body size, growth, and carcass traits have been revealed in many previous studies, including *PLAG1* (stature) (Karim et al., 2011; Bouwman et al., 2018) and *NCAPG-LCORL* (body size, carcass weight and feed efficiency) (Lindholm-Perry et al., 2011; Zhang et al., 2016; Chen et al., 2020).

Whole genome sequencing (WGS) theoretically contains all the causative mutations underlying complex traits; thus, this technology can avoid the limitation of SNP array designed by pre-selected variants and boost the ability to identify the novel candidate variants (Daetwyler et al., 2014). Recently, it is feasible and cost-effective to impute the low-density arrays to sequence level by using the small sequenced population as the reference. Imputation approach has been widely applied to detect the candidate variants for important traits in cattle (Bordbar et al., 2019; Purfield et al., 2019; Zhang et al., 2020). However, GWAS using sequencing variants may result in plenty of variants concentrating in a genomic region because of the high level of linkage disequilibrium (LD) among identified regions, which makes it difficult to accurately locate GWAS signals. Therefore, many approaches were proposed to identify the truly candidate variant from hundreds of significant variants with the high levels of LD, including condition and joint analysis (Yang et al., 2012) and Bayesian fine-mapping (Jiang et al., 2019). These strategies have been applied to precisely identify the causal variants and capture the secondary association signals for economically important traits in cattle (Fang et al., 2019; Freebern et al., 2020; Sanchez et al., 2020; Warburton et al., 2020).

Bone weight is important to body conformation and stature. Many previous studies have explored the genetic basis of bone weight using Illumina BovineHD SNP BeadChip based on different models. Several candidate genes including *LCORL*, *NCAPG*, *LAP3*, *RIMS2*, and *SLIT2* were detected for bone weight (Xia et al., 2017; Chang et al., 2018; Miao et al., 2018; Chang et al., 2019; An et al., 2020). These genes have also been identified for some other traits relevant to stature size, such as live weight, carcass weight, and body measurements (Song et al., 2016; An et al., 2018; Zhang W. et al., 2018). Despite the fact that several candidate genes and variants for bone weight have been detected in previous studies, the understanding for genetic architecture of bone weight remains to be improved with the applications of whole genome sequencing technology and advanced analysis methods.

In the current study, to explore the quantitative trait nucleotides (QTN) associated with bone weight, we firstly conducted a GWAS based on imputed whole genome sequencing variants. Using fine-mapping analysis, we further located the candidate variant within GWAS signal in the candidate region based on the Bayesian fine-mapping analysis.

MATERIALS AND METHODS

Animals and Phenotypes

A total of 1,233 Chinese Simmental beef cattle from Ulgai, Xilingol League were fed with the same conditions. A more detailed description of the feeding and management has been described previously (Zhu et al., 2017). The cattle were slaughtered at an average age of approximately 20 months. During slaughter, we measured the traits in strict accordance with the guidelines proposed by the Institutional Meat purchase Specifications for fresh beef (Zhu et al., 2017). Bone weight for

each individual was the total weight of the bones in all the forequarter and hindquarter joints. The phenotypes of bone weight were adjusted using the general linear model. Farm, year, and sex were considered as the fixed effect, and weight before fattening and fattening days were considered as covariates. Then, we considered the residuals as the pre-adjusted phenotype for the subsequent analysis.

Genotype and Imputed Sequence Variants

In total, 1,233 individuals were genotyped with the Illumina BovineHD 770 K SNP array. The SNP markers were pre-processed using PLINK v1.9 (Purcell et al., 2007; Chang et al., 2015). High-quality SNPs were selected based on the proportion of missing genotypes (<0.05), minor allele frequency (>0.01), and Hardy–Weinberg Equilibrium test ($p > 10e-6$). Also, individuals with missing genotypes (>10%) were excluded. After quality controls, a total of 641,277 autosomal SNPs were available for the subsequent imputation.

The BovineHD genotype of 1,233 individuals was imputed to whole genome sequence level based on the reference population from the unrelated 44 individuals. More detailed information about the selection of sequencing individuals and sequencing variants calling were provided in a previous study (Niu et al., 2021). We performed the imputation from SNP array to sequencing level using BEAGLE v4.1 (Browning and Browning 2007). After the imputation, a total number of 16,165,263 SNPs were obtained. The quality controls were performed including minor allele frequency (MAF) > 0.005 and imputation accuracy (R^2) > 0.1. Finally, 12,102,431 imputed sequencing SNPs were retained for the association analysis.

Genome-Wide Association Analyses

The association study for bone weight was performed using the mixed linear model in GCTA v1.93.1 software (Yang et al., 2011; Yang et al., 2014). The model is:

$$y_i = \mu + b_j x_{ij} + g_i + e_i \tag{1}$$

where y_i is the pre-adjusted phenotypic value of the ith animal, μ is the mean, b_j is the allele substitution effect of the jth SNP, x_{ij} is the jth SNP genotype of animal i and x_{ij} is coded as 0, 1, and 2 for genotypes A_1A_1, A_1A_2, and A_2A_2; g_i is the additive polygenetic effect of the ith animal, which was assumed to be distributed as $g_i \sim N(0, G\sigma_a^2)$. G is the additive genetic relationship matrix that was constructed based on all imputed SNPs using the GREML option in GCTA (Yang et al., 2010), and σ_a^2 is the genetic variance explained by all SNPs. e_i is the random residual effect, and it was assumed to be distributed as $e_i \sim N(0, I\sigma_e^2)$, where I is the identity matrix and σ_e^2 is the residual variance. The heritability for bone weight was estimated using GCTA-GREML based on BovineHD SNP array and imputed sequencing variants. The genomic heritability was computed by $h^2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_e^2}$.

The percentage of phenotypic variance and genetic variance explained by each significant SNP were calculated by $\text{Var}_p(\%) = 2pq\beta^2/\sigma_p^2 * 100\%$ and $Var_g(\%) = 2pq\beta^2/\sigma_a^2 * 100\%$, respectively, where p and q are the allele frequencies for each SNP; is the SNP allele substitution effect; σ_p^2 is the phenotypic variance; and σ_a^2 is the additive genetic variance. Here, we adopt the Bonferroni correction to

set the adjusted threshold of genome-wide significant p-values (Hayes 2013). Thus, the genome-wide significance and suggestive significance threshold were set as 9.55e-09 (p = 0.01/N) and 9.55e-07(p = 1/N), respectively, where N denotes the effective number of independent variants (a total of 1,047,272) after removing SNPs based on linkage disequilibrium (LD).

Comparison Analysis of Candidate Variants Between Previous and Current Studies

To explore evidence of the identified variants and genes from previous studies, we summarized the association results obtained by different methods based on Illumina BovineHD array including single marker association, multi-marker association, and gene-based association, and we compared their results with the finding in our study (Xia et al., 2017; Chang et al., 2018; Miao et al., 2018; Chang et al., 2019; An et al., 2020). The coordination of candidate SNPs or regions was determined according to the UMD 3.1 genome assembly.

Bayesian Fine-Mapping Analyses

To capture the independent GWAS signals in the candidate region and distinguish the possible causal variant in each signal, we performed the Bayesian fine-mapping analysis using BFMAP software (Jiang et al., 2019). The posterior probability of causality (PPC) for each variant and p-value of causality for independent association signals within candidate QTL regions were calculated to evaluate the causality of these variants. The candidate variants were determined by the causality p-value and the posterior probability in each signal. More information about BFMAP can be found at https://github.com/jiang18/bfmap.

Candidate Variants and Functional Enrichment Analyses

The candidate genes identified at suggestive threshold were located based on the UCSC database (http://genome.ucsc.edu/) in a 100-kb window (50 kb upstream and downstream of the variants). We obtained a list of candidate genes by integrating the results of current and previous studies. To explore the GO terms of candidate genes, gene set enrichment analysis was conducted using g:Profiler (Raudvere et al., 2019). The significance of GO terms was determined by the pre-adjusted g: SCS threshold (p < 0.05) in g: Profiler. The cattle QTL information was retrieved from database https://www.animalgenome.org/cgi-bin/QTLdb/BT/index (release 45, August 23, 2021). In addition, to assess the impact of SNPs on the regulation of gene expression, the information for the transcription factors (TFs) and the expression patterns of candidate genes were obtained from AnimalTFDB3.0 (http://bioinfo.life.hust.edu.cn/AnimalTFDB/) and Bgee database (https://bgee.org/).

RESULTS

GWAS Analyses for Bone Weight Using Imputed Sequence Variants

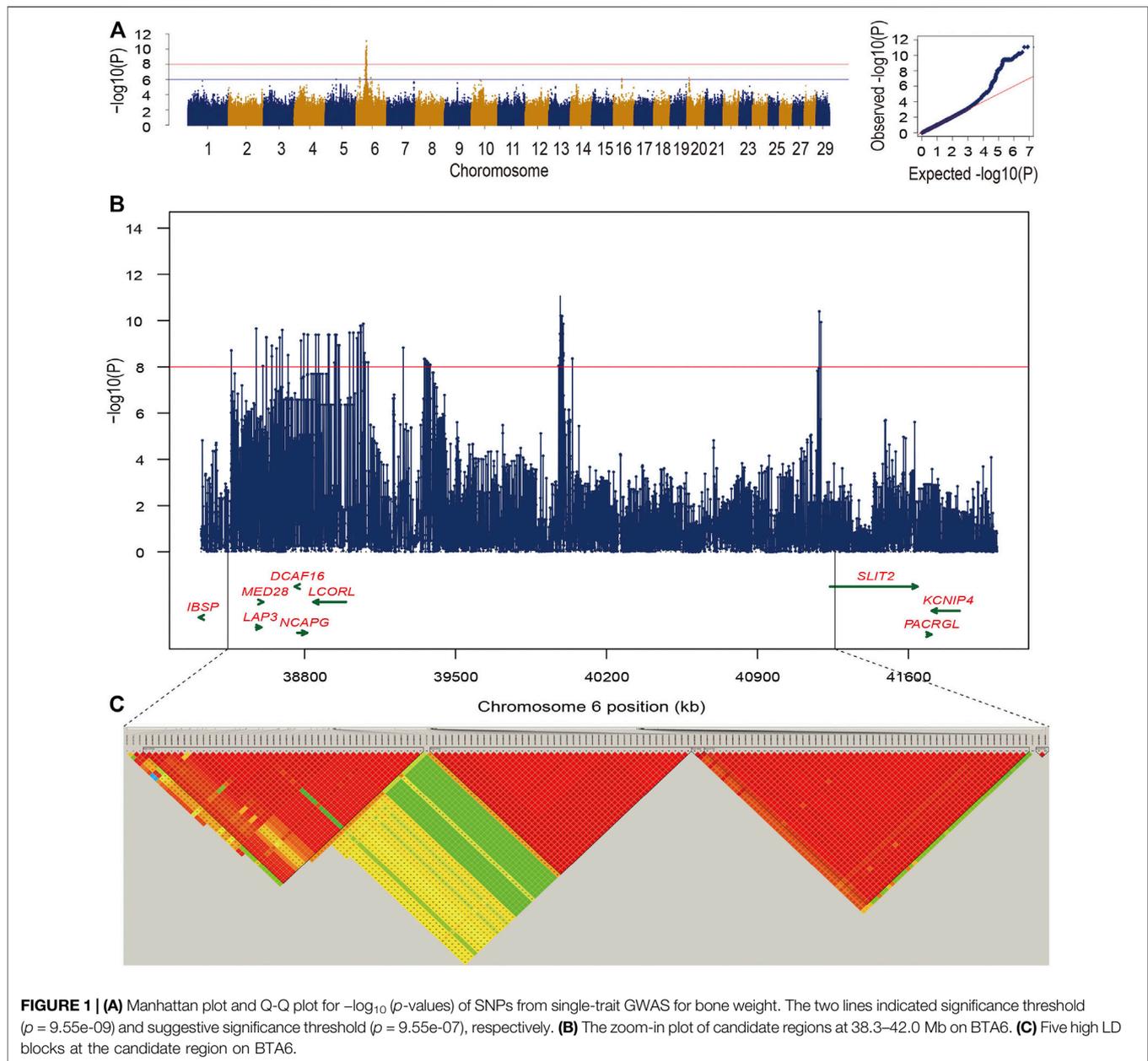In this study, we found that the heritability for bone weight estimated by Illumina BovineHD SNP array and the imputed sequence variants were 0.43 ± 0.07 and 0.44 ± 0.08, respectively. Using association analysis based on the imputed sequence variants, we totally identified 145 candidate variants for bone weight under the significant threshold. Manhattan plot presents the GWAS signals across the genome for bone weight, and the QQ-plot shows that mixed linear model is suitable for the analysis of our data (**Figure 1A**). Under the suggestive level, we found that 364 candidate variants were distributed on four chromosomes including BTA5, BTA6, BTA16, and BTA20, while most of the candidate variants were detected on BTA6.

Moreover, we found 145 significant variants located on BTA6 with the strong LD level, and five blocks with high LD were obtained in this region (**Figures 1B,C**). The most significant variant (chr6: 39989730) contributing to ~12.33% of the additive genetic variance and 5.46% of phenotypic variance is located at the intergenic region; however, no candidate gene was identified in the region at 50 kb upstream and downstream of this SNP. We also captured 10 significant intron variants within four candidate genes including *LAP3*, *MED28*, *NCAPG*, and *LCORL* (**Supplementary Table S1**). After zooming in the windows for 50 kb upstream and downstream, we observed an intergenic variant (chr6:38723514) with a MAF of 0.38, located at the 30,901 bp downstream of *DCAF16*, contributing to ~9.35% of the genetic variance and 4.14% of phenotypic variance. Meanwhile, a variant (chr6:41186810) with the p-value of 4.02e-11 at the 49,460 bp downstream of *SLIT2* can explain ~10.6% of total additive genetic variance (**Supplementary Table S1**). Under the suggestive threshold, we captured three genes containing *ABCG2*, *PKD2*, and *FAM13A* on BTA6 (**Supplementary Table S2**). An intergenic variant (p-value = 5.92e-07) was captured nearside two candidate genes including *NEUROG* and *TIFA*, while a weak LD level was observed around the 1-Mb regions (**Figure 2A**).

Based on the suggestive threshold, we identified two, one, and one candidate SNPs for bone weight on BTA5, BTA16, and BTA20 (**Supplementary Table S2**). Moreover, the suggestive variant (chr5:42704379) with a p-value of 9.07e-07 on BTA5 was located at 5,996 bp of *CPNE8*, showing a relatively strong LD with other SNPs close to *CPNE8* (**Figure 2B**). Another candidate variant on BTA16 was captured showing a strong LD with nearby SNPs, while no gene was identified nearby this region (**Figure 2C**). On BTA20, we identified a variant (chr20: 8728684) within *TNPO1* with the p-value of 5.94e−07 displaying an intense LD with nearby SNPs (**Figure 2D**).

Comparison Analyses Between Current and Previous Studies

Among 364 candidate variants, 350 candidate variants were detected by the imputed whole sequence data, while 14 variants were included in BovineHD SNP array (**Supplementary Table S2**) and 6 of them were reported in other studies. Based on these data, we identified 6 and 11 candidate genes at significant and suggestive threshold levels, respectively (**Supplementary Table S2**). We summarized the results of bone weight based on high-density SNP array using different analysis methods from many previous studies (**Supplementary Table S3**) (Xia et al., 2017; Chang et al., 2018; Miao et al., 2018; Chang et al., 2019; An et al.,

FIGURE 1 | (A) Manhattan plot and Q-Q plot for −log₁₀ (*p*-values) of SNPs from single-trait GWAS for bone weight. The two lines indicated significance threshold (*p* = 9.55e-09) and suggestive significance threshold (*p* = 9.55e-07), respectively. **(B)** The zoom-in plot of candidate regions at 38.3–42.0 Mb on BTA6. **(C)** Five high LD blocks at the candidate region on BTA6.

2020). A total of 39 candidate variants and 22 genomic regions with 35 candidate genes were obtained (Xia et al., 2017; Chang et al., 2018; Miao et al., 2018; Chang et al., 2019; An et al., 2020). Of those, seven genes were detected in the current study, including *LAP3*, *LCORL*, *NCAPG*, *SLIT2*, *FAM13A*, *MED28*, and *DCAF16*. Notably, we newly identified four genes based on the imputed whole sequence data, containing *CPNE8*, *ABCG2*, *TNPO1*, and *PKD2*.
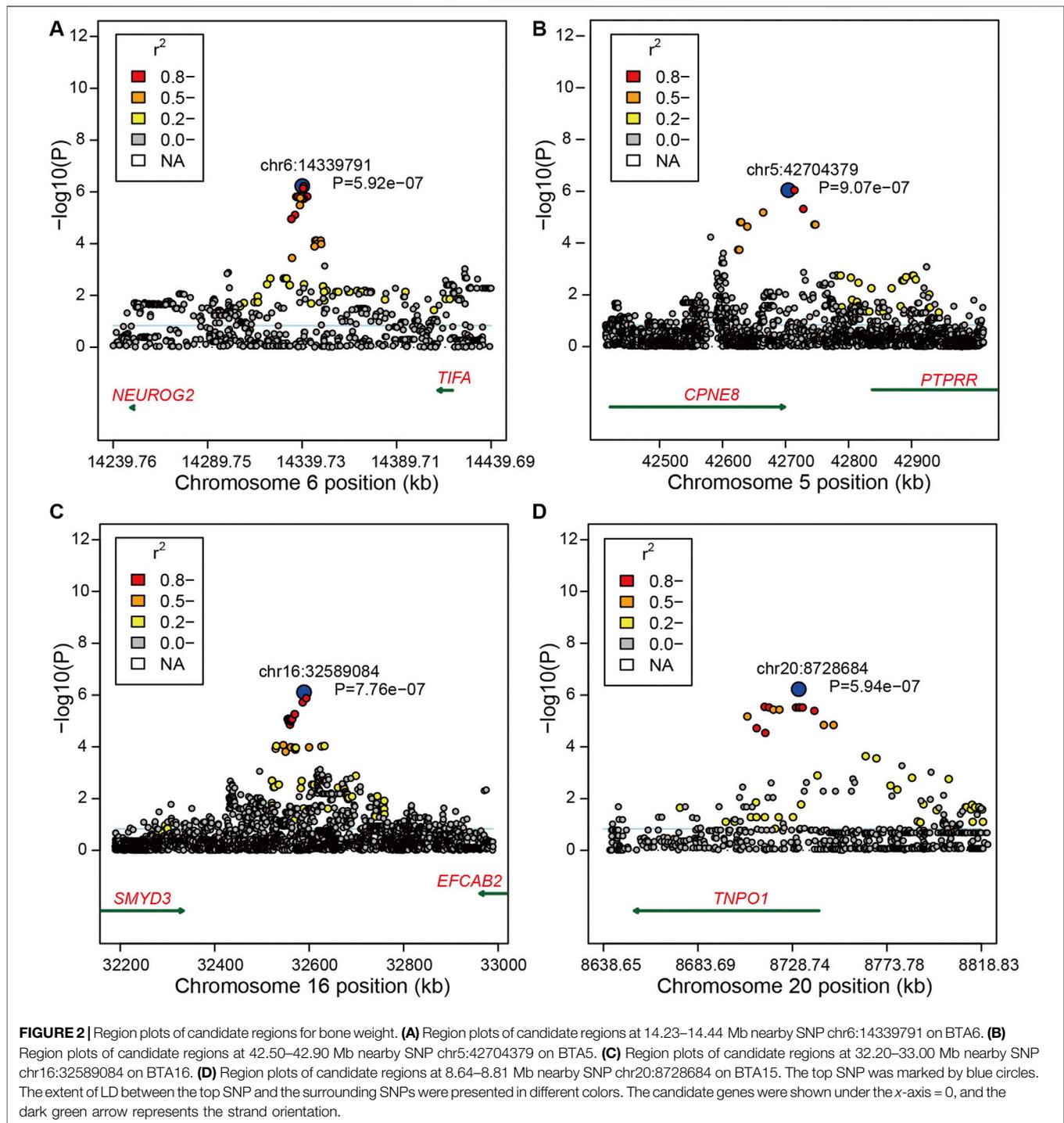
Gene Set Enrichment Analysis

We performed the enrichment analysis using 39 genes by integrating current and previous studies. These genes were enriched in eight molecular function and seven biological processes, including nucleosome-dependent ATPase activity

(GO: 0043167), cellular response to metal ion (GO: 0071248), and nervous system development (GO: 0007399). Several GO terms were related to the calcium transportation, which may be involved in bone development (**Supplementary Table S4**).

Fine-Mapping Analyses for Candidate Variants on BTA6 by Bayesian Fine-Mapping Analysis

We found that significant variants for bone weight were located on BTA6 and concentrated in a specific region (37.31–42.19 Mb) with a high LD level. To distinguish whether these variants were the significant candidates due

FIGURE 2 | Region plots of candidate regions for bone weight. **(A)** Region plots of candidate regions at 14.23–14.44 Mb nearby SNP chr6:14339791 on BTA6. **(B)** Region plots of candidate regions at 42.50–42.90 Mb nearby SNP chr5:42704379 on BTA5. **(C)** Region plots of candidate regions at 32.20–33.00 Mb nearby SNP chr16:32589084 on BTA16. **(D)** Region plots of candidate regions at 8.64–8.81 Mb nearby SNP chr20:8728684 on BTA15. The top SNP was marked by blue circles. The extent of LD between the top SNP and the surrounding SNPs were presented in different colors. The candidate genes were shown under the x-axis = 0, and the dark green arrow represents the strand orientation.

to their causality or because of their high LD with the true candidate, the Bayesian fine-mapping analysis was conducted to estimate the posterior probability of causality and causality p-value of each variant in the independent association signals. Four candidate variants including chr6:38460534, chr6: 38696766, chr6:40000601, and chr6:41194565 were identified from four independent association signals based on the Bayesian fine-mapping approach, of which the

intergenic variant (chr6:38460534) with a p-value (causality) of 1.89e-05 and a PPC of 0.6075 was identified as a potential causative variant; meanwhile, *IBSP* and *LAP3* were observed at the 137,231 bp upstream and 114,056 bp downstream of this variant. In addition, we observed another SNP (chr6:38696766) with a p-value (causality) of 5.9718e-08 and a PPC of 0.5141 located nearby *MED28*, *DCAF16*, and *NCAPG* (**Supplementary Table S5**).

DISCUSSION

The whole genome sequence variants theoretically contain more causative mutations than the SNP array, which make it more robust for the detection of quantitative trait nucleotides (Daetwyler et al., 2014). In the present study, we conducted the whole genome association studies for bone weight using the imputed sequence data, and the Bayesian fine-mapping was performed to precisely map candidate variants (Jiang et al., 2019).

Our results showed that the heritability estimated by the imputed sequencing variants is close to that by the BovineHD SNP array. Consistent with many previous studies, this finding suggested that the density of BovineHD SNP array can fully capture genetic variance, while the sequence variants may provide redundant information due to LD (van Binsbergen et al., 2015; Veerkamp et al., 2016; Zhang C. et al., 2018). Another explanation is that low imputation accuracy may lead to heritability missing estimated by imputed WGS data (Xavier et al., 2016; Song et al., 2019).

Identification of candidate variants for bone weight in beef cattle have been extensively conducted using multiple strategies based on BovineHD SNP array. For instance, using the gene-based association analysis and single marker association study, a previous study revealed 11 potential variants and several important candidate genes affecting bone weight, such as *FAM184B*, *LAP3*, and *NCAPG* (Xia et al., 2017). Another study used the three association models including fixed polygene, random polygene, and composite interval mapping polygene, and identified a total of seven unique variants and five candidate genes associated with bone weight (*C12ORF74*, *LCORL*, *RIMS2*, *WDFY3*, and *FER1L6*) (Chang et al., 2018; Miao et al., 2018; Chang et al., 2019). Moreover, other methods based on LASSO and LMM model including multi-marker association and bin model have also been applied to obtain several important candidate genes for bone weight in previous studies (*RIMS2*, *LCORL*, and *SLIT2*) (An et al., 2020). However, GWAS in previous studies may be biased due to the design of the SNP array; thus, we conducted the GWAS using the imputed whole genome sequencing data and Bayesian fine-mapping to capture more variants for bone weight.

In the present study, we detected a total of 364 candidate variants for bone weight, 350 of them were not included in the BovineHD SNP array. We annotated 11 potential genes including *SLIT2*, *LAP3*, *MED28*, *NCAPG*, *LCORL*, *DCAF16*, *CPNE8*, *FAM13A*, *ABCG2*, *PKD2*, and *TNPO1*. *SLIT2* embedded with two intergenic variants was captured for bone weight using high-density arrays; this gene was reported to be associated with growth and carcass traits including spleen weight (An et al., 2018), birth weight (Smith et al., 2019), bone weight (Xia et al., 2017; An et al., 2020), carcass weight, and eye muscle area (Bhuiyan et al., 2018) in cattle. *LAP3*, *MED28*, *NCAPG*, *LCORL*, and *DCAF16* were identified within the significant genomic region (chr6: 37315342–42194093), captured by 24 candidate variants. *LAP3*, as a member of LAPs family involved in cell maintenance and growth development, had been suggested that it has potential function on body weight in

sheep (La et al., 2019) and body measurements traits in cattle (An et al., 2019). Moreover, previous studies also reported that *LAP3* is the potential candidate gene for feed efficiency component traits (Lindholm-Perry et al., 2011; Zhang et al., 2016), carcass merit, and internal organ traits including kidney weight and spleen weight (Song et al., 2016; Xia et al., 2017; An et al., 2018; Miao et al., 2018). *MED28*, a gene involved in the regulation of cell proliferation and cycle (Cho et al., 2019), was captured to associate with body weight, intramuscular fat content, and yearling weight in cattle (Santiago et al., 2017a; Santiago et al., 2017b; Anton et al., 2018). Additionally, two significant intron variants are located within *NCAPG*, which is a well-known gene involved in cell proliferation (Seipold et al., 2009) and has been reported to be associated with feed intake, average daily gain (Lindholm-Perry et al., 2011; Zhang et al., 2016), and carcass traits including hot carcass weight and lean meat yield (Wang et al., 2020). *LCORL* and *DCAF16* were also identified for bone weight in our study, and these two genes have been captured as the candidate genes for calving ease (Bongiorni et al., 2012; Sahana et al., 2015), carcass and meat quality traits, as well as feed efficiency component traits in cattle, which may be directly related to the stature size (Zhang et al., 2016; Anton et al., 2018; Wang et al., 2020).

To evaluate the potential impact of the candidate variants on the gene expression, we queried the TFs existing in the six candidate genes that were located in the BTA6 block based on the AnimalTFDB3.0. Notably, we identified a TF cofactor (Cofactors Family: Mediator complex) related to *MED28* and a TF related to *LCORL* (TF Family: HTH) in *Bos taurus*, which displayed high ortholog identity and conservation across multiple species. Moreover, using the gene expression data obtained from the Bgee database, we observed that *MED28* has a high expression level in prefrontal cortex, brain, etc., but a relatively low expression level in skeletal muscle tissue, and *LCORL* was also observed expressing in testis, colon, brain, and skeletal muscle tissue. In addition, we observed a total of 8,697 regions overlapped with the cattle QTLdb (release 45, August 23, 2021), involved with many economic traits related to the stature size, including average daily gain, body weight, bone weight, and body measurements traits.

Under the suggestive threshold, a total of eight variants were identified nearby *FAM13A* on BTA6; *FAM13A* was captured to be associated with bone weight using a gene-based association method (Xia et al., 2017). Compared with many previous studies, we identified four candidate genes, namely, *CPNE8*, *ABCG2*, *PKD2*, and *TNP O 1*. Of those, *CPNE8* and *ABCG2* were firstly reported for bone weight in this study. *PKD2* was identified by an intron variant (chr6:38024322) with a *p*-value of 4.52e-08, which was related to calcium homeostasis and involved in mitotic cell cycle process (Olsen et al., 2007; Abo-Ismail et al., 2014). *PKD2* was detected as one of the candidate genes for birth weight, weaning weight, and yearling weight in cattle (Weng et al., 2016). Remarkably, *PKD2* was identified nearby the candidate QTLs for bone percentage, meat percentage, and meat-to-bone ratio in previous studies (Gutiérrez-Gil et al., 2009; Abo-Ismail et al., 2014). *TNPO1* was embedded with an intergenic variant (chr20:8728684) with a *p*-value of 5.94e-07; this

gene was detected to be associated with birth weight, mature weight, and yearling weight in cattle, which implied its role in regulation of the stature in cattle (Weng et al., 2016).

Using the Bayesian fine-mapping approach, we narrowed the scope of candidate region and identified the putative variants on BTA6. A total of four independent signals were identified in the candidate region, and seven genes nearby the putative variants were captured. Notably, *IBSP* was observed at 137,231 bp upstream of an intergenic variant (chr6:38460534) with posterior probability 0.85; this gene has been reported to relate to the structure of bone matrix in human (Denninger et al., 2015) and body weight in cattle (Santiago et al., 2017a). Moreover, many previous studies suggested that *IBSP* has an impact on the skeletal development in mouse (Rivadeneira et al., 2009), sheep (Matika et al., 2016), and cattle (Gutiérrez-Gil et al., 2012; Xu et al., 2018).

CONCLUSION

Using imputed sequencing variants, we identified a total of 350 variants that were not included in BovineHD SNP array in Chinese Simmental beef cattle. Our study revealed several candidate genes and GO terms for bone weight. Moreover, we identified four potential variants on BTA6 by the Bayesian fine-mapping approach. Our findings provide insights into understanding the genetic basis of bone weight that may affect stature size, and these results could be potentially applied in a breeding program in cattle.

DATA AVAILABILITY STATEMENT

The data underlying this study have been uploaded to Dryad. The raw genotype data are accessible using the following https://doi.org/10.5061/dryad.4qc06.

ETHICS STATEMENT

The animal study was reviewed and approved by the Institute of Animal Sciences, Chinese Academy of Agricultural Sciences.

AUTHOR CONTRIBUTIONS

LX, JL, and QN conceived and designed the experiments. QN and LX analyzed the data. LX, TZ, TW, ZW, XG, YC, BZ, and HG contributed reagents/materials/analysis tools. QN wrote the paper. All authors read and approved the final manuscript.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.750746/full#supplementary-material

REFERENCES

Abo-Ismail, M. K., Vander Voort, G., Squires, J. J., Swanson, K. C., Mandell, I. B., Liao, X., et al. (2014). Single Nucleotide Polymorphisms for Feed Efficiency and Performance in Crossbred Beef Cattle. *BMC Genet.* 15, 14. doi:10.1186/1471-2156-15-14

Albertí, P., Panea, B., Sañudo, C., Olleta, J. L., Ripoll, G., Ertbjerg, P., et al. (2008). Live Weight, Body Size and Carcass Characteristics of Young Bulls of Fifteen European Breeds. *Livestock Sci.* 114, 19–30. doi:10.1016/j.livsci.2007.04.010

An, B., Gao, X., Chang, T., Xia, J., Wang, X., Miao, J., et al. (2020). Genome-wide Association Studies Using Binned Genotypes. *Heredity* 124, 288–298. doi:10.1038/s41437-019-0279-y

An, B., Xia, J., Chang, T., Wang, X., Miao, J., Xu, L., et al. (2018). Genome-wide Association Study Identifies Loci and Candidate Genes for Internal Organ Weights in Simmental Beef Cattle. *Physiol. Genomics* 50, 523–531. doi:10.1152/physiolgenomics.00022.2018

An, B., Xia, J., Chang, T., Wang, X., Xu, L., Zhang, L., et al. (2019). Genome-wide Association Study Reveals Candidate Genes Associated with Body Measurement Traits in Chinese Wagyu Beef Cattle. *Anim. Genet.* 50, 386–390. doi:10.1111/age.12805

Anton, I., Húth, B., Füller, I., Rózsa, L., Holló, G., and Zsolnai, A. (2018). Effect of Single Nucleotide Polymorphisms on Intramuscular Fat Content in Hungarian Simmental Cattle. *Asian-australas J. Anim. Sci.* 31, 1415–1419. doi:10.5713/ajas.17.0773

Berg, R. T., and Butterfield, R. M. (1966). Muscle: Bone Ratio and Fat Percentage as Measures of Beef Carcass Composition. *Anim. Sci.* 8, 1–11. doi:10.1017/s000335610003765x

Bhuiyan, M. S. A., Lim, D., Park, M., Lee, S., Kim, Y., Gondro, C., et al. (2018). Functional Partitioning of Genomic Variance and Genome-wide Association Study for Carcass Traits in Korean Hanwoo Cattle Using Imputed Sequence Level SNP Data. *Front. Genet.* 9, 217. doi:10.3389/fgene.2018.00217

Bongiorni, S., Mancini, G., Chillemi, G., Pariset, L., and Valentini, A. (2012). Identification of a Short Region on Chromosome 6 Affecting Direct Calving Ease in Piedmontese Cattle Breed. *PLoS One* 7, e50137. doi:10.1371/journal.pone.0050137

Bordbar, F., Jensen, J., Zhu, B., Wang, Z., Xu, L., Chang, T., et al. (2019). Identification of Muscle-specific Candidate Genes in Simmental Beef Cattle Using Imputed Next Generation Sequencing. *PLoS One* 14, e0223671. doi:10.1371/journal.pone.0223671

Bouwman, A. C., Daetwyler, H. D., Chamberlain, A. J., Ponce, C. H., Sargolzaei, M., Schenkel, F. S., et al. (2018). Meta-analysis of Genome-wide Association Studies for Cattle Stature Identifies Common Genes that Regulate Body Size in Mammals. *Nat. Genet.* 50, 362–367. doi:10.1038/s41588-018-0056-5

Browning, B. L., and Browning, S. R. (2009). A Unified Approach to Genotype Imputation and Haplotype-phase Inference for Large Data Sets of Trios and

Unrelated Individuals. *Am. J. Hum. Genet.* 84, 210–223. doi:10.1016/
j.ajhg.2009.01.005

Browning, S. R., and Browning, B. L. (2007). Rapid and Accurate Haplotype
Phasing and Missing-Data Inference for Whole-Genome Association Studies
by Use of Localized Haplotype Clustering. *Am. J. Hum. Genet.* 81, 1084–1097.
doi:10.1086/521987

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., and Lee, J. J.
(2015). Second-generation PLINK: Rising to the challenge of Larger and Richer
Datasets. *GigaSci* 4, 7. doi:10.1186/s13742-015-0047-8

Chang, T., Wei, J., Liang, M., An, B., Wang, X., Zhu, B., et al. (2019). A Fast and
Powerful Empirical Bayes Method for Genome-wide Association Studies.
Animals (Basel) 9, 305. doi:10.3390/ani9060305

Chang, T., Xia, J., Xu, L., Wang, X., Zhu, B., Zhang, L., et al. (2018). A Genome-
wide Association Study Suggests Several Novel Candidate Genes for Carcass
Traits in Chinese Simmental Beef Cattle. *Anim. Genet.* 49, 312–316.
doi:10.1111/age.12667

Chen, Q., Huang, B., Zhan, J., Wang, J., Qu, K., Zhang, F., et al. (2020). Whole-
genome Analyses Identify Loci and Selective Signals Associated with Body Size
in Cattle. *J. Anim. Sci.* 98, skaa068. doi:10.1093/jas/skaa068

Cho, J. G., Choi, J. S., Lee, J. H., Cho, M. G., Kim, H. S., Noh, H. D., et al. (2019).
MED28 Over-expression Shortens the Cell Cycle and Induces Genomic
Instability. *Int. J. Mol. Sci.* 20, 1746. doi:10.3390/ijms20071746

Chumlea, W. C., Wisemandle, W., Guo, S. S., and Siervogel, R. M. (2002). Relations
between Frame Size and Body Composition and Bone mineral Status. *Am.
J. Clin. Nutr.* 75, 1012–1016. doi:10.1093/ajcn/75.6.1012

Conroy, S. B., Drennan, M. J., McGee, M., Keane, M. G., Kenny, D. A., and Berry,
D. P. (2010). Predicting Beef Carcass Meat, Fat and Bone Proportions from
Carcass Conformation and Fat Scores or Hindquarter Dissection. *Animal* 4,
234–241. doi:10.1017/s1751731109991121

Daetwyler, H. D., Capitan, A., Pausch, H., Stothard, P., van Binsbergen, R.,
Brøndum, R. F., et al. (2014). Whole-genome Sequencing of 234 Bulls
Facilitates Mapping of Monogenic and Complex Traits in Cattle. *Nat.
Genet.* 46, 858–865. doi:10.1038/ng.3034

Denninger, K. C., Litman, T., Marstrand, T., Moller, K., Svensson, L., Labuda, T.,
et al. (2015). Kinetics of Gene Expression and Bone Remodelling in the Clinical
Phase of Collagen-Induced Arthritis. *Arthritis Res. Ther.* 17, 43. doi:10.1186/
s13075-015-0531-7

Fang, L., Jiang, J., Li, B., Zhou, Y., Freebern, E., Vanraden, P. M., et al. (2019).
Genetic and Epigenetic Architecture of Paternal Origin Contribute to Gestation
Length in Cattle. *Commun. Biol.* 2, 100. doi:10.1038/s42003-019-0341-6

Fang, Z.-H., and Pausch, H. (2019). Multi-trait Meta-Analyses Reveal 25
Quantitative Trait Loci for Economically Important Traits in Brown Swiss
Cattle. *BMC Genomics* 20, 695. doi:10.1186/s12864-019-6066-6

Freebern, E., Santos, D. J. A., Fang, L., Jiang, J., Parker Gaddis, K. L., Liu, G. E., et al.
(2020). GWAS and fine-mapping of Livability and Six Disease Traits in
Holstein Cattle. *BMC genomics* 21, 41. doi:10.1186/s12864-020-6461-z

Gutiérrez-Gil, B., Williams, J. L., Homer, D., Burton, D., Haley, C. S., and Wiener,
P. (2009). Search for Quantitative Trait Loci Affecting Growth and Carcass
Traits in a Cross Population of Beef and Dairy Cattle. *J. Anim. Sci.* 87, 24–36.
doi:10.2527/jas.2008-0922

Gutiérrez-Gil, B., Wiener, P., Williams, J. L., and Haley, C. S. (2012). Investigation
of the Genetic Architecture of a Bone Carcass Weight QTL on BTA6. *Anim.
Genet.* 43, 654–661. doi:10.1111/j.1365-2052.2012.02322.x

Hayes, B. (2013). Overview of Statistical Methods for Genome-wide Association Studies
(GWAS). *Methods Mol. Biol.* 1019, 149–169. doi:10.1007/978-1-62703-447-0_6

Jiang, J., Cole, J. B., Freebern, E., Da, Y., VanRaden, P. M., and Ma, L. (2019).
Functional Annotation and Bayesian fine-mapping Reveals Candidate Genes
for Important Agronomic Traits in Holstein Bulls. *Commun. Biol.* 2, 212.
doi:10.1038/s42003-019-0454-y

Karim, L., Takeda, H., Lin, L., Druet, T., Arias, J. A. C., Baurain, D., et al. (2011).
Variants Modulating the Expression of a Chromosome Domain Encompassing
PLAG1 Influence Bovine Stature. *Nat. Genet.* 43, 405–413. doi:10.1038/ng.814

La, Y., Zhang, X., Li, F., Zhang, D., Li, C., Mo, F., et al. (2019). Molecular
Characterization and Expression of SPP1, LAP3 and LCORL and Their
Association with Growth Traits in Sheep. *Genes (Basel)* 10, 616.
doi:10.3390/genes10080616

Lindholm-Perry, A. K., Sexten, A. K., Kuehn, L. A., Smith, T. P., King, D. A.,
Shackelford, S. D., et al. (2011). Association, Effects and Validation of

Polymorphisms within the NCAPG - LCORL Locus Located on BTA6 with
Feed Intake, Gain, Meat and Carcass Traits in Beef Cattle. *BMC Genet.* 12, 103.
doi:10.1186/1471-2156-12-103

Mader, C. J., Montanholi, Y. R., Wang, Y. J., Miller, S. P., Mandell, I. B., McBride, B.
W., et al. (2009). Relationships Among Measures of Growth Performance and
Efficiency with Carcass Traits, Visceral Organ Mass, and Pancreatic Digestive
Enzymes in Feedlot Cattle1,2. *J. Anim. Sci.* 87, 1548–1557. doi:10.2527/jas.2008-
0914

Matika, O., Riggio, V., Anselme-Moizan, M., Law, A. S., Pong-Wong, R., Archibald,
A. L., et al. (2016). Genome-wide Association Reveals QTL for Growth, Bone
and *In Vivo* Carcass Traits as Assessed by Computed Tomography in Scottish
Blackface Lambs. *Genet. Sel Evol.* 48, 11. doi:10.1186/s12711-016-0191-3

Miao, J., Wang, X., Bao, J., Jin, S., Chang, T., Xia, J., et al. (2018). Multimarker and
Rare Variants Genomewide Association Studies for Bone Weight in Simmental
Cattle. *J. Anim. Breed. Genet.* 135, 159–169. doi:10.1111/jbg.12326

Niu, Q., Zhang, T., Xu, L., Wang, T., Wang, Z., Zhu, B., et al. (2021). Integration of
Selection Signatures and Multi-Trait GWAS Reveals Polygenic Genetic
Architecture of Carcass Traits in Beef Cattle. *Genomics* 113, 3325–3336.
doi:10.1016/j.ygeno.2021.07.025

Olsen, H. G., Nilsen, H., Hayes, B., Berg, P. R., Svendsen, M., Lien, S., et al. (2007).
Genetic Support for a Quantitative Trait Nucleotide in the ABCG2 Gene
Affecting Milk Composition of Dairy Cattle. *BMC Genet.* 8, 32. doi:10.1186/
1471-2156-8-32

Ozkaya, S., and Bozkurt, Y. (2008). The Relationship of Parameters of Body
Measures and Body Weight by Using Digital Image Analysis in Pre-slaughter
Cattle. *Arch. Anim. Breed.* 51, 120–128. doi:10.5194/aab-51-120-2008

Pabiou, T., Fikse, W. F., Amer, P. R., Cromie, A. R., Näsholm, A., and Berry, D. P.
(2012). Genetic Relationships between Carcass Cut Weights Predicted from
Video Image Analysis and Other Performance Traits in Cattle. *Animal* 6,
1389–1397. doi:10.1017/s1751731112000705

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D.,
et al. (2007). PLINK: a Tool Set for Whole-Genome Association and
Population-Based Linkage Analyses. *Am. J. Hum. Genet.* 81, 559–575.
doi:10.1086/519795

Purfield, D. C., Evans, R. D., and Berry, D. P. (2019). Reaffirmation of Known
Major Genes and the Identification of Novel Candidate Genes Associated with
Carcass-Related Metrics Based on Whole Genome Sequence within a Large
Multi-Breed Cattle Population. *BMC Genomics* 20, 720. doi:10.1186/s12864-
019-6071-9

Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., et al. (2019).
g:Profiler: a Web Server for Functional Enrichment Analysis and Conversions
of Gene Lists (2019 Update). *Nucleic Acids Res.* 47, W191–w198. doi:10.1093/
nar/gkz369

Rivadeneira, F., Styrkársdottir, U., Estrada, K., Halldórsson, B. V., Hsu, Y. H.,
Richards, J. B., et al. (2009). Twenty bone-mineral-density Loci Identified by
Large-Scale Meta-Analysis of Genome-wide Association Studies. *Nat. Genet.*
41, 1199–1206. doi:10.1038/ng.446

Sahana, G., Höglund, J. K., Guldbrandtsen, B., and Lund, M. S. (2015). Loci
Associated with Adult Stature Also Affect Calf Birth Survival in Cattle. *BMC
Genet.* 16, 47. doi:10.1186/s12863-015-0202-3

Sanchez, M.-P., Guatteo, R., Davergne, A., Saout, J., Grohs, C., Deloche, M.-C., et al.
(2020). Identification of the ABCC4, IER3, and CBFA2T2 Candidate Genes for
Resistance to Paratuberculosis from Sequence-Based GWAS in Holstein and
Normande Dairy Cattle. *Genet. Sel Evol.* 52, 14. doi:10.1186/s12711-020-00535-9

Santiago, G. G., Siqueira, F., Cardoso, F. F., Regitano, L. C. A., Ventura, R., Sollero, B. P.,
et al. (2017a). Genomewide Association Study for Production and Meat Quality
Traits in Canchim Beef Cattle. *J. Anim. Sci.* 95, 3381–3390. doi:10.2527/jas2017.1570

Santiago, G. G., Siqueira, F., Cardoso, F. F., Regitano, L. C. A., Ventura, R., Sollero,
B. P., et al. (2017b). Genomewide Association Study for Production and Meat
Quality Traits in Canchim Beef Cattle1. *J. Anim. Sci.* 95, 3381–3390.
doi:10.2527/jas.2017.1570

Seipold, S., Priller, F. C., Goldsmith, P., Harris, W. A., Baier, H., and Abdelilah-
Seyfried, S. (2009). Non-SMC Condensin I Complex Proteins Control
Chromosome Segregation and Survival of Proliferating Cells in the
Zebrafish Neural Retina. *BMC Dev. Biol.* 9, 40. doi:10.1186/1471-213x-9-40

Sieber, M., Freeman, A. E., and Kelley, D. H. (1988). Relationships between Body
Measurements, Body Weight, and Productivity in Holstein Dairy Cows. *J. Dairy
Sci.* 71, 3437–3445. doi:10.3168/jds.s0022-0302(88)79949-x

Smith, J. L., Wilson, M. L., Nilson, S. M., Rowan, T. N., Oldeschulte, D. L., Schnabel, R. D., et al. (2019). Genome-wide Association and Genotype by Environment Interactions for Growth Traits in U.S. Gelbvieh Cattle. *BMC Genomics* 20, 926. doi:10.1186/s12864-019-6231-y

Snowder, G. D., Van Vleck, L. D., Cundiff, L. V., Bennett, G. L., Koohmaraie, M., and Dikeman, M. E. (2007). Bovine Respiratory Disease in Feedlot Cattle: Phenotypic, Environmental, and Genetic Correlations with Growth, Carcass, and Longissimus Muscle Palatability Traits1. *J. Anim. Sci.* 85, 1885–1892. doi:10.2527/jas.2007-0008

Song, H., Ye, S., Jiang, Y., Zhang, Z., Zhang, Q., and Ding, X. (2019). Using Imputation-Based Whole-Genome Sequencing Data to Improve the Accuracy of Genomic Prediction for Combined Populations in Pigs. *Genet. Sel Evol.* 51, 58. doi:10.1186/s12711-019-0500-8

Song, Y., Xu, L., Chen, Y., Zhang, L., Gao, H., Zhu, B., et al. (2016). Genome-Wide Association Study Reveals the PLAG1 Gene for Knuckle, Biceps and Shank Weight in Simmental Beef Cattle. *PLoS One* 11, e0168316. doi:10.1371/journal.pone.0168316

van Binsbergen, R., Calus, M. P. L., Bink, M. C. A. M., van Eeuwijk, F. A., Schrooten, C., and Veerkamp, R. F. (2015). Genomic Prediction Using Imputed Whole-Genome Sequence Data in Holstein Friesian Cattle. *Genet. Sel Evol.* 47, 71. doi:10.1186/s12711-015-0149-x

Veerkamp, R. F., Bouwman, A. C., Schrooten, C., and Calus, M. P. L. (2016). Genomic Prediction Using Preselected DNA Variants from a GWAS with Whole-Genome Sequence Data in Holstein-Friesian Cattle. *Genet. Sel Evol.* 48, 95. doi:10.1186/s12711-016-0274-1

Wang, Y., Zhang, F., Mukiibi, R., Chen, L., Vinsky, M., Plastow, G., et al. (2020). Genetic Architecture of Quantitative Traits in Beef Cattle Revealed by Genome Wide Association Studies of Imputed Whole Genome Sequence Variants: II: Carcass merit Traits. *BMC Genomics* 21, 38. doi:10.1186/s12864-019-6273-1

Warburton, C. L., Engle, B. N., Ross, E. M., Costilla, R., Moore, S. S., Corbet, N. J., et al. (2020). Use of Whole-Genome Sequence Data and Novel Genomic Selection Strategies to Improve Selection for Age at Puberty in Tropically-Adapted Beef Heifers. *Genet. Sel Evol.* 52, 28. doi:10.1186/s12711-020-00547-5

Weng, Z., Su, H., Saatchi, M., Lee, J., Thomas, M. G., Dunkelberger, J. R., et al. (2016). Genome-wide Association Study of Growth and Body Composition Traits in Brangus Beef Cattle. *Livestock Sci.* 183, 4–11. doi:10.1016/j.livsci.2015.11.011

Wu, X., Fang, M., Liu, L., Wang, S., Liu, J., Ding, X., et al. (2013). Genome Wide Association Studies for Body Conformation Traits in the Chinese Holstein Cattle Population. *BMC Genomics* 14, 897. doi:10.1186/1471-2164-14-897

Xavier, A., Muir, W. M., and Rainey, K. M. (2016). Impact of Imputation Methods on the Amount of Genetic Variation Captured by a Single-Nucleotide Polymorphism Panel in Soybeans. *BMC bioinformatics* 17, 55. doi:10.1186/s12859-016-0899-7

Xia, J., Fan, H., Chang, T., Xu, L., Zhang, W., Song, Y., et al. (2017). Searching for New Loci and Candidate Genes for Economically Important Traits through Gene-Based Association Analysis of Simmental Cattle. *Sci. Rep.* 7, 42048. doi:10.1038/srep42048

Xu, L., Zhang, W. G., Shen, H. X., Zhang, Y., Zhao, Y. M., Jia, Y. T., et al. (2018). Genome-wide Scanning Reveals Genetic Diversity and Signatures of Selection in Chinese Indigenous Cattle Breeds. *Livestock Sci.* 216, 100–108. doi:10.1016/j.livsci.2018.08.005

Yang, J., Ferreira, T., Morris, A. P., Medland, S. E., Madden, P. A., Heath, A. C., et al. (2012). Conditional and Joint Multiple-SNP Analysis of GWAS Summary Statistics Identifies Additional Variants Influencing Complex Traits. *Nat. Genet.* 44, 369–375. doi:10.1038/ng.2213

Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., et al. (2010). Common SNPs Explain a Large Proportion of the Heritability for Human Height. *Nat. Genet.* 42, 565–569. doi:10.1038/ng.608

Yang, J., Lee, S. H., Goddard, M. E., and Visscher, P. M. (2011). GCTA: a Tool for Genome-wide Complex Trait Analysis. *Am. J. Hum. Genet.* 88, 76–82. doi:10.1016/j.ajhg.2010.11.011

Yang, J., Zaitlen, N. A., Goddard, M. E., Visscher, P. M., and Price, A. L. (2014). Advantages and Pitfalls in the Application of Mixed-Model Association Methods. *Nat. Genet.* 46, 100–106. doi:10.1038/ng.2876

Zhang, C., Kemp, R. A., Stothard, P., Wang, Z., Boddicker, N., Krivushin, K., et al. (2018a). Genomic Evaluation of Feed Efficiency Component Traits in Duroc Pigs Using 80K, 650K and Whole-Genome Sequence Variants. *Genet. Sel Evol.* 50, 14. doi:10.1186/s12711-018-0387-9

Zhang, F., Wang, Y., Mukiibi, R., Chen, L., Vinsky, M., Plastow, G., et al. (2020). Genetic Architecture of Quantitative Traits in Beef Cattle Revealed by Genome Wide Association Studies of Imputed Whole Genome Sequence Variants: I: Feed Efficiency and Component Traits. *BMC Genomics* 21, 36. doi:10.1186/s12864-019-6362-1

Zhang, W., Gao, X., Shi, X., Zhu, B., Wang, Z., Gao, H., et al. (2018b). PCA-based Multiple-Trait GWAS Analysis: A Powerful Model for Exploring Pleiotropy. *Animals (Basel)* 8, 239. doi:10.3390/ani8120239

Zhang, W., Li, J., Guo, Y., Zhang, L., Xu, L., Gao, X., et al. (2016). Multi-strategy Genome-wide Association Studies Identify the DCAF16-NCAPG Region as a Susceptibility Locus for Average Daily Gain in Cattle. *Sci. Rep.* 6, 38073. doi:10.1038/srep38073

Zhu, B., Niu, H., Zhang, W., Wang, Z., Liang, Y., Guan, L., et al. (2017). Genome Wide Association Study and Genomic Prediction for Fatty Acid Composition in Chinese Simmental Beef Cattle Using High Density SNP Array. *BMC Genomics* 18, 464. doi:10.1186/s12864-017-3847-7