

## RESEARCH ARTICLE

# Identification of muscle-specific candidate genes in Simmental beef cattle using imputed next generation sequencing

Farhad Bordbar<sup>1</sup>, Just Jensen<sup>2</sup>, Bo Zhu<sup>1</sup>, Zezhao Wang<sup>1</sup>, Lei Xu<sup>1</sup>, Tianpeng Chang<sup>1</sup>, Ling Xu<sup>1</sup>, Min Du<sup>3</sup>, Lupei Zhang<sup>1</sup>, Huijiang Gao<sup>1</sup>, Lingyang Xu<sup>1\*</sup>, Junya Li<sup>1\*</sup>

**1** Laboratory of Molecular Biology and Bovine Breeding, Institute of Animal Science, Chinese Academy of Agricultural Sciences, Beijing, China, **2** Department of Molecular Biology and Genetics, Aarhus University, Aarhus, Denmark, **3** Department of Animal Sciences, Washington Center for Muscle Biology, Washington State University, Pullman, Washington, United States of America

\* [lijunya@caas.cn](mailto:lijunya@caas.cn) (JYL); [xulingyang@caas.cn](mailto:xulingyang@caas.cn) (LYX)



## OPEN ACCESS

**Citation:** 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(10): e0223671. <https://doi.org/10.1371/journal.pone.0223671>

**Editor:** Julio Cesar de Souza, Universidade Federal de Mato Grosso do Sul, BRAZIL

**Received:** April 8, 2019

**Accepted:** September 25, 2019

**Published:** October 10, 2019

**Copyright:** © 2019 Bordbar et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** The data underlying this study have been uploaded to Dryad. The raw genotype data are accessible using the following doi: [10.5061/dryad.4qc06](https://doi.org/10.5061/dryad.4qc06). Additional data are available using the doi: [10.5061/dryad.5v3k1ct](https://doi.org/10.5061/dryad.5v3k1ct).

**Funding:** This study was supported by the National Natural Science Foundations of China (31372294 to JL), Chinese Academy of Agricultural Sciences of Technology Innovation Project (CAAS-ZDXT2018006 and ASTIP-IAS03), Project of College Innovation Improvement under Beijing

## Abstract

Genome-wide association studies (GWAS) have commonly been used to identify candidate genes that control economically important traits in livestock. Our objective was to detect potential candidate genes associated mainly with muscle development traits related to dimension of hindquarter in cattle. A next generation sequencing (NGS) dataset to imputed to 12 million single nucleotide polymorphisms (SNPs) (from 1252 Simmental beef cattle) were used to search for genes affecting hindquarter traits using a linear, mixed model approach. We also used haplotype and linkage disequilibrium blocks to further support our identifications. We identified 202 significant SNPs in the bovine BTA4 chromosome region associated with width of hind leg, based on a stringent statistical threshold ( $p = 0.05/\text{effective number of SNPs identified}$ ). After exploring the region around these SNPs, we found candidate genes that were potentially related to the associated markers. More importantly, we identified a region of approximately 280 Kb on the BTA4 chromosome that harbored several muscle specific candidate genes, genes to be in a potential region for muscle development. However, we also found candidate gene *SLC13A1* on BTA4, which seems to be associated with bone disorders (such as chondrodysplasia) in Simmental beef cattle.

## Introduction

Cattle and other livestock species are economically important in almost all regions of the world [1]. Cattle breeds are continuously being developed to optimize the efficiency of meat production [2] and meat-production traits, controlled by a variety of genes, are among the most important traits selected for in various beef cattle breeding programs [3]. More specifically, cattle breeders desire cattle with more rapid lean muscle growth. The focus on lean muscle is important because consumers today demand leaner meat [4], thus making muscle growth an important economic factor in the beef industry. Therefore, within the past few years, more research has focused on economically important traits affecting the quality of

Municipality (PXM2016\_014207\_000012) and national beef cattle industrial technology system (CARS-37). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

meat production in livestock, particularly on muscle development. However, breeding programs are having difficulty addressing muscle quality problems in livestock, such as poor meat quality in swine, in part due to insufficient information about underlying mechanisms involved with muscle growth and its metabolism [5].

A candidate gene approach, used in genetic association studies for a variety of species [6], might be useful for improving our understanding of genes responsible for regulating traits involved with the underlying metabolism of muscle growth. That is, because candidate genes possess specific biological actions relative to various aspects of a target trait, the candidate gene approach can provide a valuable strategy for detecting the QTL (quantitative trait loci) that controls genetic variation of an important trait [7].

Genome-wide association studies (GWAS) provide a robust approach for identifying genetic variants that exert prominent effects on quantitative traits and complex diseases in humans [8–10] and farm animals [11]. Recently, several association studies related to economically important traits (production, health, reproduction and body conformation) in cattle have been published [12–14]. However, the success of the GWAS approach depends on a variety of factors, such as sample size and the number of genetic variants affecting the trait [15].

Next generation sequencing (NGS), a recently developed method, has been used to incorporate more genetic variants than SNP (single nucleotide polymorphism) arrays. Generally, an NGS dataset exhibits a higher error rate compared to genotyping based on DNA chips, however, this problem can be remedied by imputing a larger number of genetic variants. Applying NGS in association studies offers numerous advantages over traditional DNA sequencing approaches, including higher throughput sequencing, more specific applicability, and higher read quality for sequencing dataset [16]. Using an NGS dataset, Sharma et al. pinpointed 18 mutations associated with Mendelian diseases in Hanwoo cattle [17], whereas 33 genes were identified as being associated with the domestication process [18]. The NGS approach is also useful for investigating genes associated with complex traits. For example, three loci have been identified in the European domestic pigs [19], including the *PLAG1* and *LCORL* genes, both of which are associated with growth rate and meat quality in many livestock species. However, one of the problems associated with using an NGS dataset is capacity of typical computers to calculate marker effects, due to high density of SNPs [20].

Previous studies have found that Chromosomes 2, 3, 6, 14, 20, and 29 exert significant effects (location for various novel QTLs) on important carcass traits such as meat production in cattle [21–25]. In fact, muscle specific genes are distributed over several chromosomes. For example, *Alpha-actin* (*ACTA1*) is located on bovine Chromosome 28 (identified using various analyses [26–27]), whereas bovine Chromosome 5 harbors *Myogenic factor 5* (*MYF5*), first identified by linkage mapping [28]. Using linkage analysis, Charlier et al. identified the locus for double muscling associated with the *MH* gene on bovine Chromosome 2 [29]. The important muscle-specific gene *myogenin* (*MYOG*) was mapped on bovine Chromosome 16 [30]. *Calpain 1* (*CAPN1*) and *calpastatin* (*CAST*) genes, which play important roles in meat tenderness, have been mapped on bovine Chromosomes 29 and 7, respectively [31, 32].

Hind quarter provide the most valuable cuts on the beef carcass and can, therefore, be advantageous specifically for growth and meat leanness traits. The hind quarter include connective tissues, major weight supporting bones such as femur, tibia, and fibula and variety of covering muscle tissues, such as gastrocnemius, biceps femoris, semitendinosus, semimembranosus, vastus lateralis, tensor fascia lata, and gluteus medius. Almost all of these muscles are large skeletal muscles that tend to produce copious amounts of tender meat. Piedrafita et al. reported a close correlation between the dimension of the hind leg and percent muscle [33]. Moreover, a direct association was documented between thickness of hind limb and growth rate of European beef cattle [33].

Many papers have focused on hind legs relative to muscle development, bone disorders, and protein metabolism [34–39]. In this study, we use the GWAS approach to identify candidate genes associated with the width, length, and circumference of hind legs. More specifically, the major aim of this research was to identify markers and QTLs that significantly affect the size of the hind quarter in the Simmental breed. We suspect that various factors could potentially affect lean meat production in the hind leg, hind-leg bone characteristics, and location of fat in the hind leg. We expect our result to elucidate the underlying molecular mechanism involved with hind-leg meat production in Simmental cattle. To our knowledge, our study is the first to apply GWAS analysis to attempt to identify candidate genes responsible for hind leg related traits in Simmental beef cattle.

## Materials and methods

### Ethics statement

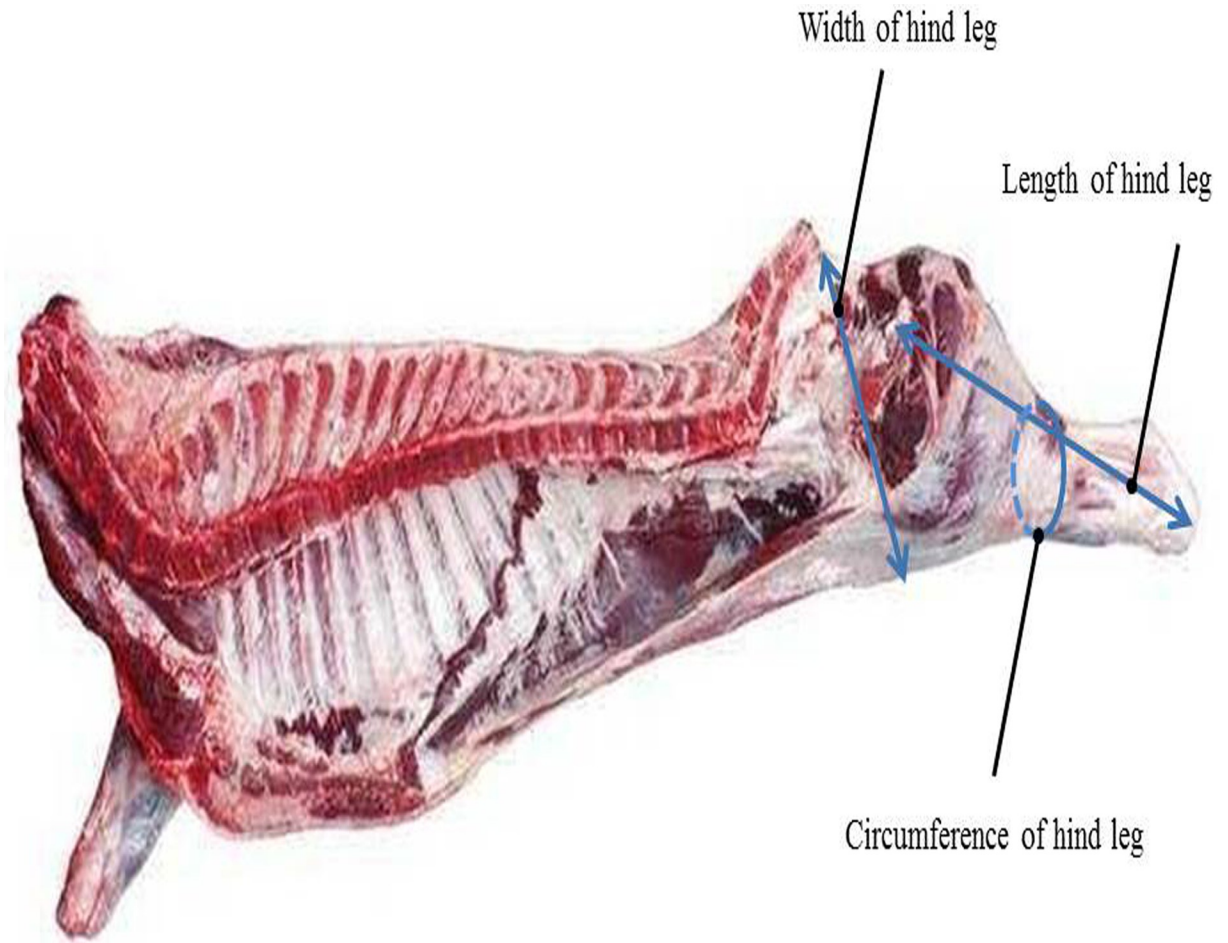
We followed guidelines advanced by the China Council on Animal Care. All steps and activities were authorized by the Institute of Animal Science (IAS) and the Chinese Academy of Agricultural Science (CAAS), Beijing, China. Animal examinations followed rules of the China's Animal Welfare council.

### Animal resources and phenotype data

Our population included 1346 Simmental beef cattle born from 2009 to 2015. The animals originated from the Chinese regions of Ulgai, Xilingol League, and Inner Mongolia. After weaning, calves were transferred to the Beijing Jinweifuren farm feedlot. All calves were raised under the same feeding and management protocols. At 16 to 18 months of age, animals were slaughtered via electric stunning. [Due to the long distance from the feed lot to the abattoir (almost 332 km, 10-h drive), all cattle were rested for almost one day before slaughtering them to make sure that they had recovered from traveling and any related stress.] After bloodletting, carcasses were split (cut open) vertically into two equal parts and then weighed. Split carcasses were immediately placed in refrigerators at 4°C until examined traits could be measured (approximately nine hours after slaughtering). Examined traits were measured according to strict guidelines outlined by Institutional Meat Purchase Specifications for fresh beef. Measurements of the hind leg included the width, length, and circumference (Fig 1). Width of hind leg is the straight-line distance from the inside of the tail recess to the leading edge of the thigh (<https://www.greatfoodsolutions.com/products/beef-carcass-regular-cut>. Accessed 23 NOV. 2018), length of hind leg is the distance from the anterior border of the pubic suture to the midpoint of the ankle joint, and circumference of the hind leg is horizontal circumference at the junction of the femur and tibia. Mean, standard deviation, maximum, and minimum of the measured traits are listed in Table 1.

### Genotype data and quality control of SNP array

We contracted with the Illumina Genome Studio (San Diego, CA) to sample genotypes and analyze SNP chips. To quality control, we removed cattle with high Pi-Hat values and subjected the other cattle to exclusion criteria consisting of SNPs with call rates less than 99%, minor allele frequencies (MAF) less than 5%, and severe deviations from Hardy-Weinberg equilibrium ( $P < 10^{-6}$ ). In addition, individuals with >10% missing genotypes were removed from the dataset. All these criteria were processed by PLINK v1.07 [40]. After control quality, 1252 Simmental beef cattle with 671,204 autosomal SNPs remained for subsequent analyses.



**Fig 1. Locations of width, length, and circumference of hind leg measurements for the experimental beef carcasses.**

<https://doi.org/10.1371/journal.pone.0223671.g001>

## Resequencing

Based on genomic relationships and Pi-Hat values, we chose 44 unrelated Simmental beef cattle for genetic resequencing of whole genome. We used the TIANamp Blood DNA kit (Tiangen Biotech Company Limited, Beijing, China) to construct NGS library, which consisted of choosing DNA with an A260/280 nm absorbance ratio ranging between 1.8 and 2.0. We prepared one library for each animal ( $n = 44$ ) using an Illumina Hiseq 2500 genome sequencing system (Illumina Inc., San Diego, CA, USA). We obtained 9,621,765,847 reads, which we then subjected to our quality control process. We excluded low-quality reads, such as those containing more than 10% unknown bases, more than 10% mismatches, or more than 50% low-quality bases. We also removed putative PCR duplicates produced by PCR amplification in the library construction process.

**Table 1. Descriptive statistics for hind-leg measurements of experimental animals.**

Trait	Mean (cm)	Standard deviation (cm)	Maximum (cm)	Minimum (cm)
The width of hind leg	44.86	4.13	67	25
The length of hind leg	76.78	5.16	92	50
The circumference of hind leg	68.01	16.58	110	22

<https://doi.org/10.1371/journal.pone.0223671.t001>

We used 1.5 µg of DNA per sample as input material for each sample. We generated sequencing libraries using the Truseq Nano DNA HT sample preparation Kit (Illumina Inc., San Diego, CA, USA) following manufacturer's recommendations and added index codes to attribute sequences for each sample. We fragmented each DNA sample with sonication to 350 bp, then DNA fragments were end polished, A-tails added, and ligated with the full-length adapter for Illumina sequencing with further PCR amplification (Fig 2). We purified PCR products with an AMPure XP system and analyzed library size distributions with an Agilent 2100 Bioanalyzer and then quantified them using real-time PCR. After implementing our quality control procedures, we obtained 9,584,920,309 reads. Sample reads and quality scores (Q), which represents the quality of sequencing data [41], are listed in S1 Table. The average sequencing depth for each sample was about 20X with a range of 17X to 25X (X represents a number of times which each base is sequenced).

### Imputation of sequence variants

Sequence SNPs with MAF more than 0.05 were subjected to imputation. Based on 21,043,398 sequence variants in the 44 genetically sequenced cattle, we performed imputation using BEAGLE v4.1 [42] with its default parameter settings. Beagle is based on an algorithm using population-based information to infer haplotypes and missing genotypes. We coded genotypes for imputed sequence variants as 0 (for homozygotes), 1 (for heterozygotes), or 2 (for alternative homozygotes). R-sq. data (showing imputation quality [43]) larger than 0.1 were retained for association studies. Following imputation, we obtained 12,468,401 markers for Chromosomes 1–29.

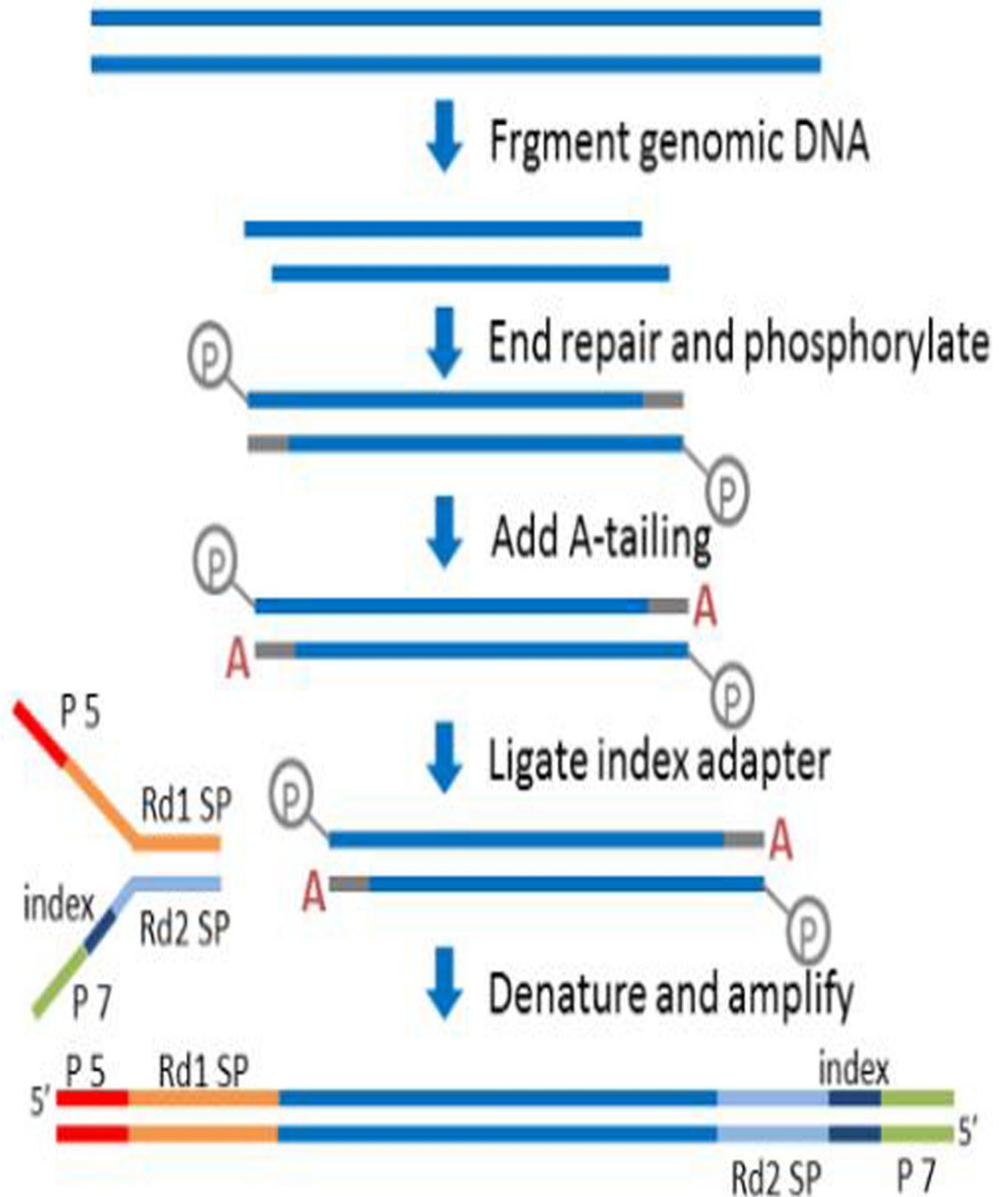
### Statistical analysis

We used a general mixed linear model, determined as follows,

$$y = \mu + Xb_i + mjb_j + Zu + e$$

Where,  $y$  is the vector of phenotypic values for traits of interest and  $\mu$  is the population mean. We used single marker regression with two types of fixed effect variables, including  $b_p$ , representing the vector corresponding to the nuisance fixed effect (including birth year, sex, weight, number of fattening days) and  $b_j$ , the SNP effect and  $m_j$  representing the vectors of the  $i^{\text{th}}$  marker. Variable  $u$  is the polygenic effect with the assumption of  $N(0, \sigma^2 K)$ , where  $K$  is defined as the kinship matrix (all markers on autosomal chromosomes were used except for those markers which were under investigation in  $b_j$ ) and  $\sigma^2$  is the additive genetic variance. Variable  $X$  is the incidence matrix relating the phenotypic observations to fixed effects, whereas  $Z$  is the matrix relating the phenotypic observations to polygenic effects. Variable  $e$  is a vector for random residual effects with the assumption that  $V(e) = I\sigma_e^2$ , where  $I$  represents the identity matrix and  $\sigma_e^2$  represents residual variance.

We used the GenABEL v1.8–0 [44] algorithm in the R package to identify the statistical significance of each SNP; based Bonferroni corrected P-values of 0.05 divided by the effective number of SNPs. We calculated the t statistic to determine the P-values of each SNP. We used the UCSC genome browser (<http://www.genome.ucsc.edu>) to search for candidate genes. We estimated haplotype blocks by PLINK v1.07 [41] using default commands (*plink—bfile mydata—blocks*) which calculated LD (linkage disequilibrium) for SNPs within 200 kb in length. To check the association of each haplotype in block, following parameters were used: *plink—bfile mydata—hap plink.blocks—hap-freq* [41]. The Haploview v4.2 algorithm [45] was applied to detect haplotype blocks illustrated for each chromosome.



**Fig 2. Library construction plot.** The plot was provided by Genome Sequencing company Novogene (<https://en.novogene.com>). <https://doi.org/10.1371/journal.pone.0223671.g002>

### SNP distribution

One problem associated with analyzing NGS data is the high throughput density of SNPs, which makes it difficult to compute p-values of the SNPs with computers typically accessible by researchers [20]. To address this problem, we divided the bovine chromosomes into three segments: Segment 1 consisting of Chromosomes 1–10 (5,830,727 SNPs), Segment 2 consisting of Chromosomes 11–20 (4,063,690 SNPs), and Segment 3 consisting of Chromosomes 21–

29 (2,573,984 SNPs). We applied different p-value thresholds for each segment due to the different numbers of SNPs in each segment.

## Results

### Association analysis

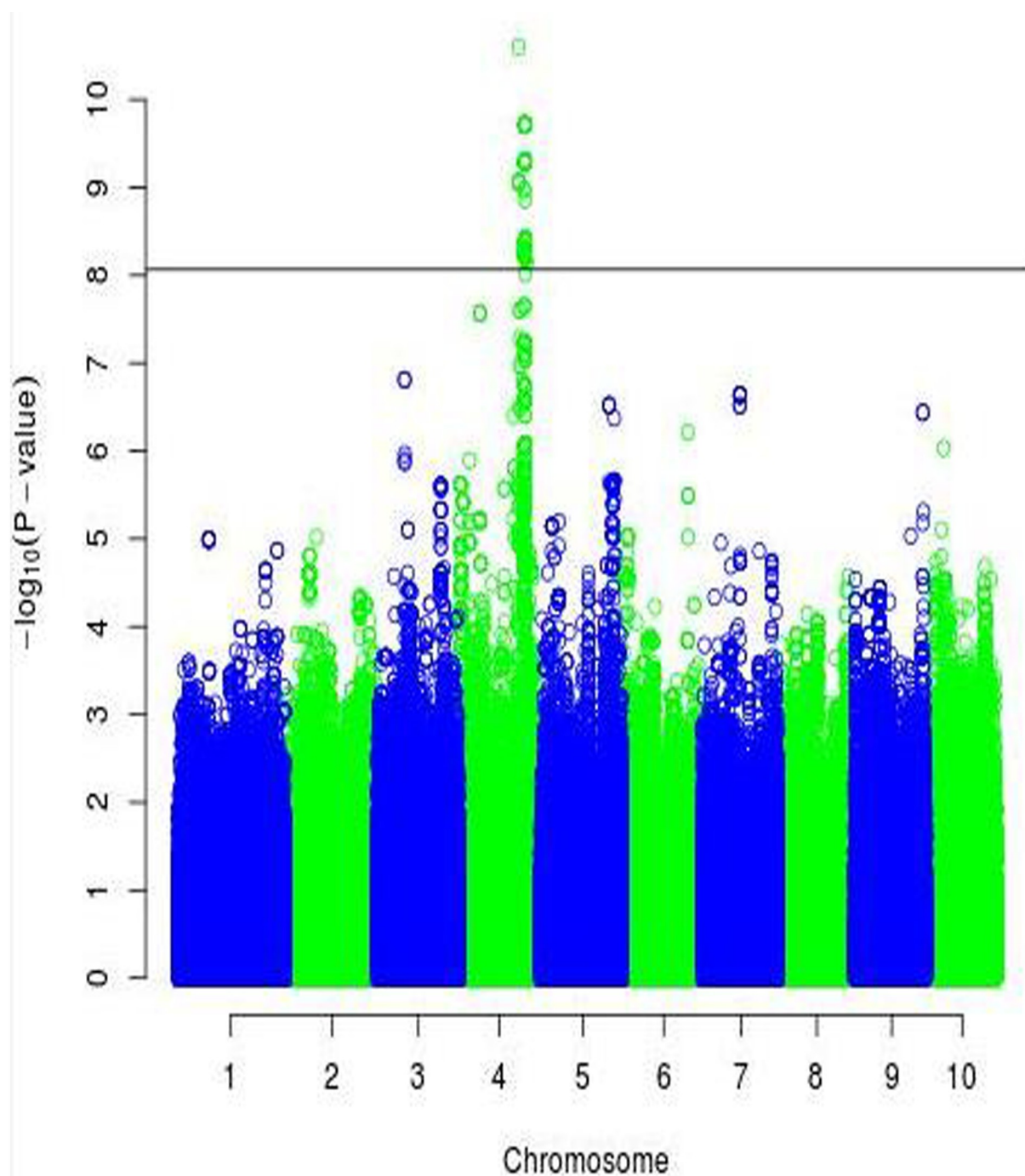
To identify more reliable QTLs and candidate genes, we set a stringent threshold to each segment ( $p = 0.05/\text{effective number of SNPs}$ ) [i.e., Segment 1 ( $p = 8.58 \times 10^{-9}$ ), Segment 2 ( $p = 1.24 \times 10^{-8}$ ), and Segment 3 ( $p = 1.95 \times 10^{-8}$ )]. Fig 3 illustrates the Manhattan plot for the significance of Segment 1 (associated with the width of the hind leg). Surprisingly, we identified 202 SNPs in the BTA4 chromosome that surpassed our established significance threshold, thus making them potential candidates for QTLs (related to the width of the hind leg). We divided these significant SNPs into Sub-sections 1–5 based on their physical locations: (1) 88 Mb (7 SNPs), (2) 96 Mb (18 SNPs), (3) 97 Mb (115 SNPs), (4) 98 Mb (61 SNPs), and (5) 100 Mb (1 SNP). We identified three candidate genes (*IQUB*, *NDUFA5*, and *ASB15*) for Sub-section 1, which spanned with a range of about 322 kb. These three genes were located in close vicinity to one another, totally separated by only 30 kb. Of the seven associated SNPs in Sub-section 1, two SNPs (*chr4:88718818* and *chr4:88724320*) were located in the *ASB15* gene region, and two SNPs (*chr4:88581497* and *chr4:88619410*) were located in the *IQUB* gene region. Moreover, by expanding the region about 45 kb downstream of the highest associated SNP with width of hind leg (*chr4:88402194*), we found candidate gene *SLC13A1*. We also identified two more candidate genes, *LMOD2* and *WASL*, approximately 50 kb upstream of gene *ASB15*. Haplotype analysis showed both flanking SNPs in Sub-section 1 were located in two haplotype blocks, each spanning 27 kb. Fig 4 illustrates Linkage disequilibrium plot for the most significant SNP (*chr4:88402194*) ( $P = 2.52 \times 10^{-11}$ ) located in high LD block.

For Sub-section 2, we obtained one candidate gene (*PLXNA4*) with 18 significant SNPs. The distance between these 18 SNPs was about 130 kb and interestingly, all these SNPs were located within the *PLXNA4* gene (464 kb long). Most SNPs were located in or near high LD blocks, such as SNP *chr4:96893352*, located in the 3-kb-high LD block shown in Fig 5.

Sub-section 3 corresponded to 115 associated SNPs spanning about 90 kb. After we noticed there was no candidate gene between the flanking SNPs of *chr4:97748185* ( $p = 1.93 \times 10^{-10}$ ) and *chr4:97657432* ( $p = 6.16 \times 10^{-9}$ ), we investigated 100 kb downstream and upstream of both flanking SNPs. For SNP *chr4:97748185*, we discovered the candidate gene *EXOC4*, located at almost 50 kb upstream of the SNP. Moreover, for SNP *chr4:97657432*, we identified the candidate gene *CHCHD3* about 50 kb downstream of the SNP. Although there was no candidate gene between the above two markers, we found about 10 haplotype blocks totally spanning 45 kb, which suggests the importance of the 115 markers identified in this region of the genome.

The *EXOC4* gene was also located in the Sub-section 4, which spanned around 26 kb of BTA4 chromosome. This candidate gene is one of the largest genes, with an approximate length of 800 kb. More importantly, the *EXOC4* gene harbored all significant SNPs in Sub-section 4 (61 SNPs) and most of these SNPs which totally corresponded to 32 kb haplotype blocks were near or within high LD blocks. In addition, after searching for more candidate genes, we noticed the location of the micro-RNA *MIR2423* (73 bp), which was almost 55 kb upstream from SNP *chr4:98014505*.

There was only one SNP (*chr4:100895756*) associated with fifth sub-section of Segment 1. When investigating the region around this SNP, we identified a very potentially important candidate gene for Simmental beef cattle, called *MTPN*, located approximately 250 kb downstream from the SNP. (We are in the process of validating this gene). The width of statistically significant hind leg SNPs are listed in S2 Table.

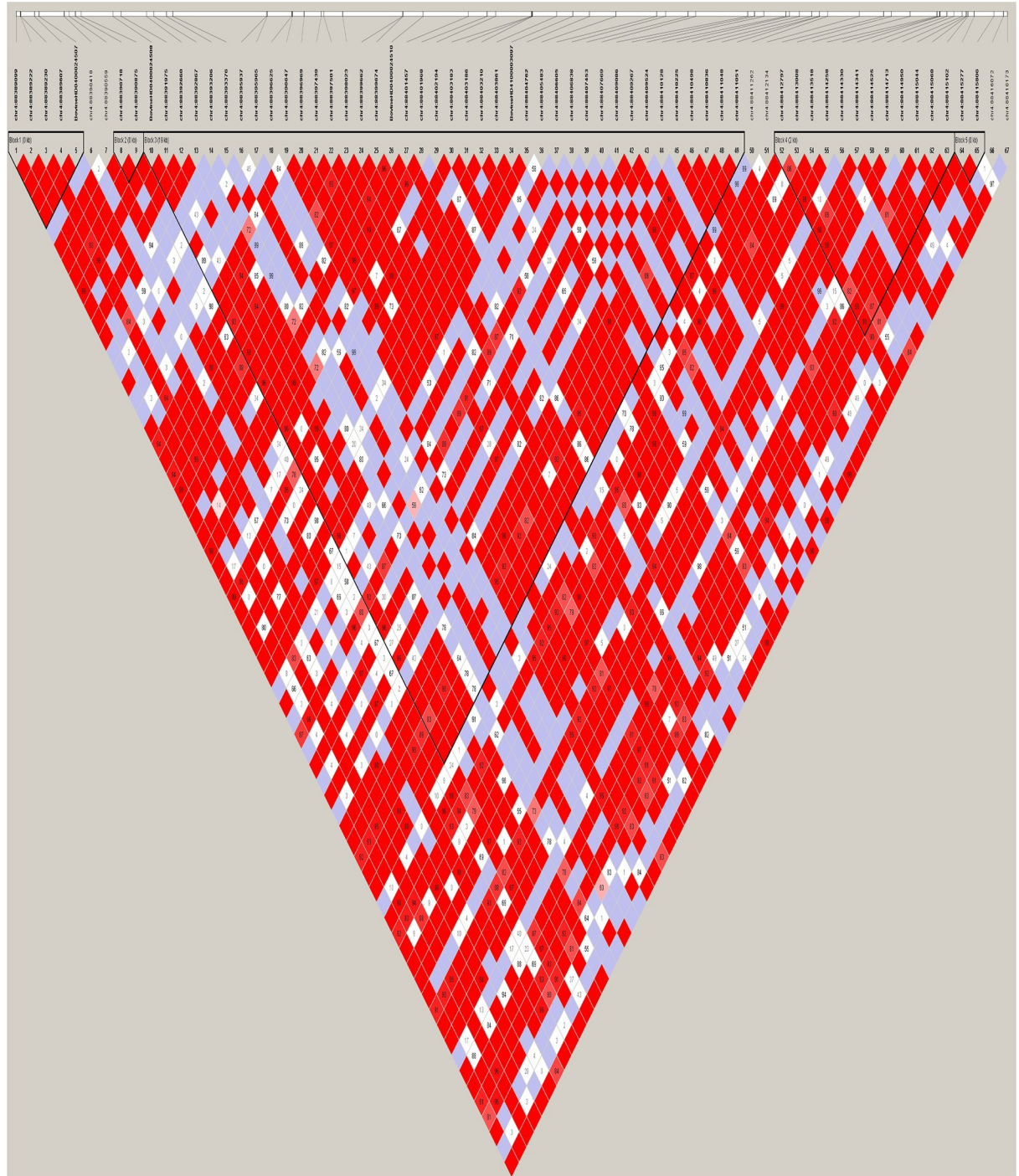


**Fig 3. Manhattan plot of  $-\log_{10}(p)$  for the region associated with the width of hind leg.** The horizontal line represents the assigned threshold of significance.

<https://doi.org/10.1371/journal.pone.0223671.g003>

We found only one SNP (*chr5:91327124*) in BTA5 chromosome associated with the circumference of hind leg, ( $p = 1.23 \times 10^{-9}$ ). After exploring the region around the SNP, we did not find any candidate genes in close proximity. Even when we tolerated a less stringent threshold, we could not identify any significant SNPs associated with hind leg length.



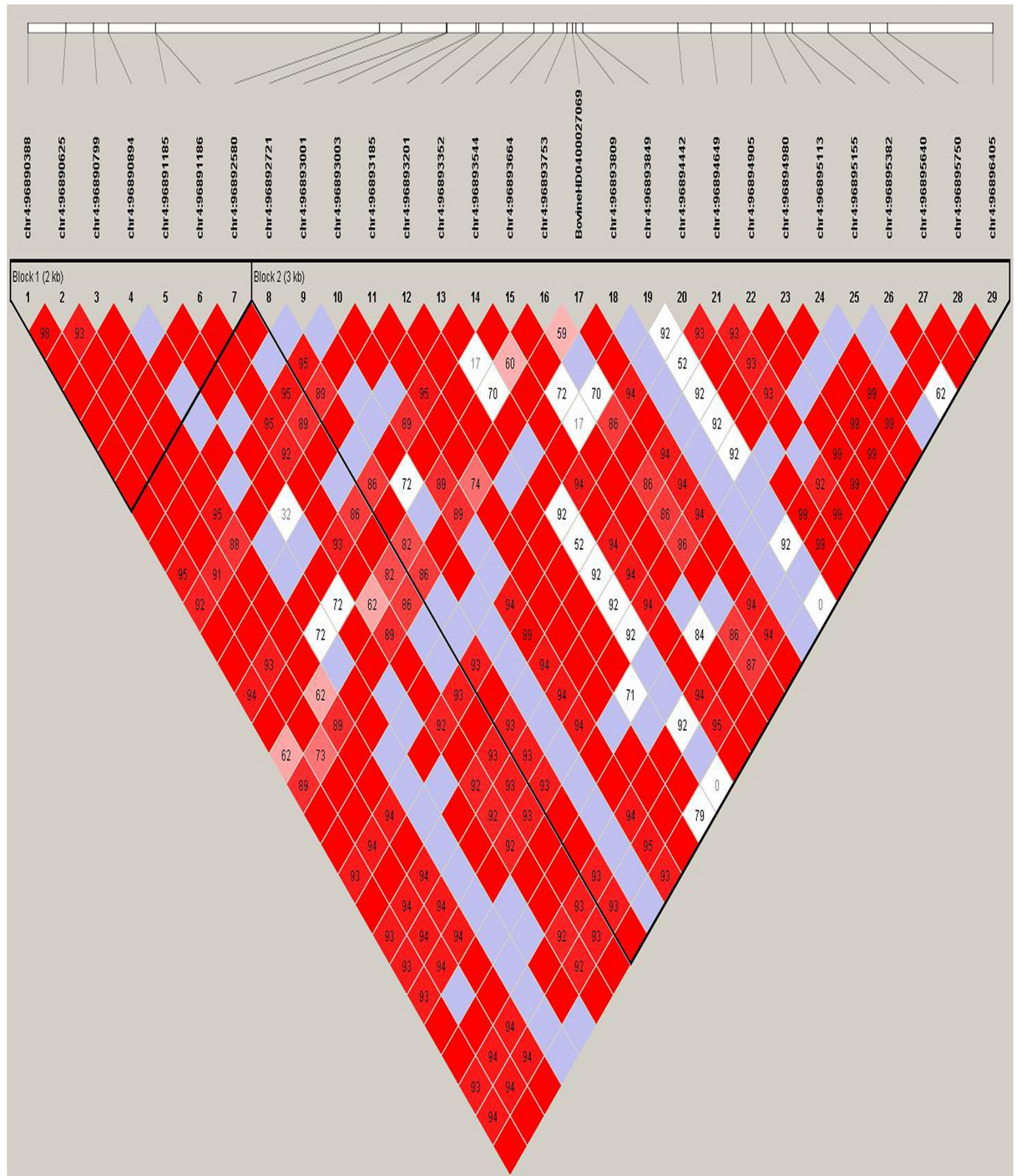


**Fig 4. Linkage disequilibrium plot.** The high LD block spans 19 kb, which harbors the most significant SNP for width of hind leg.

<https://doi.org/10.1371/journal.pone.0223671.g004>

**Discussion**

Genome-wide association studies (GWAS) have been applied widely for plants [46] and animals [24]. This method has produced immense strides in elucidating molecular mechanism underlying commercially desirable traits in livestock, such as in beef and dairy cattle



**Fig 5. Linkage disequilibrium plot.** The plot reveals a high LD block spanning 3 kb around the associated SNP *chr4:96893352* in Sub-section 2.

<https://doi.org/10.1371/journal.pone.0223671.g005>

[14, 47–52]. In other words, by facilitating the identification of candidate genes with GWAS, breeders are more able to address problems associated with breeding programs (i.e., selecting animals with higher production) and improve desirable traits. Another advantage of GWAS is that entire genomes can be investigated for complex traits examining haplotype

blocks and linkage disequilibrium data among genetic markers. Therefore, GWAS is a powerful method for identifying quantitative trait loci and regions that harbor causative genes [53–55]. We expect that the next generation sequencing approach will rapidly improve resequencing procedures for identifying enormous numbers of SNPs, which will in turn enable researchers to identify rare genetic variants and more intensely investigate basic and important mechanisms underlying genetic traits [16].

Interestingly, among the candidate genes located in the BTA4 chromosome, most seem to be muscle-related. *ASB15* is a conserved gene expressed mostly in the musculoskeletal system. The gene *ASB15* is involved in a variety of cell processes, such as proliferation and differentiation [56–58]. *ASB15* also plays a major role in signal transduction pathways encouraged by  $\beta$ -adrenergic agonists (BA) receptor compounds in hypertrophied skeletal muscles [59].

McDanel et al. found that the *ASB15* gene induces protein turnover and synthesis and so stimulates myoblast differentiation and muscle cell development [60]. In addition, McDanel and Spurlock reported that *ASB15* alters C2C12 mouse myoblast differentiation via mediation of the PI3K/Akt signal transduction pathway [61]. *NDUFA5* is expressed in a variety of tissues, including in skeletal muscle [62]. This gene, which is involved in transportation of electrons in mitochondria, and is also associated with autism in humans [63]. When conducting miRNA microarray analysis, Seong et al. showed an association between *NDUFA5* and marbling score in Hanwoo cattle [64]. When the amount of feed for Holstein-Friesian bulls was restricted, expression of the *NDUFA5* gene increased and energy production was enhanced [65].

Several studies have proven the existence of *LMOD2* mRNAs in striated muscle [66–68]. In addition, knocking out the *LMOD2* gene can shorten thin filament (actin) prior to dilated cardiomyopathy [69]. Similarly, Yuen et al. showed that the *LMOD2* gene is involved in the regulating the length of thin filaments [70]. After gene expression validation, the *LMOD2* gene has also been shown to be related with muscle development in cattle [71].

In our study, the genes *ASB15*, *NDUFA5*, and *LMOD2* were surrounded by *IQUB* and *WASL* genes (both in close proximity). Furthermore, it has been shown in a cancer research that knockdown of *IQUB* gene can prevent *c-myc* expression [72] (*c-myc* is a muscle regulator gene which is involved with proliferation and differentiation of muscle myoblast).

We also noticed the association of the *WASL* gene which act in the cytoskeleton [73], which is a complex network of cells playing a role in muscle contraction. Taken together, due to all mentioned genes above are related with muscle (either connected with muscle structural proteins or muscle development) we hypothesize that the chromosome region on BTA4 (88 Mb spanning about 280 kb), which supports five closely located genes, is a candidate region for skeletal muscle development genes in Simmental beef cattle. Although the functions of some of the genes occurring in this region have been identified, more research is needed to verify the identity of other candidate genes, specifically regarding their association with differentiation and proliferation of skeletal muscle cells.

Humans and several other species, including *Bos taurus*, can contract a serious genetic disease called chondrodysplasia (dwarfism) which causes abnormality in bone shape and skeleton structure [74]. A mutation (deletion) in gene *SLC4A2* is related to osteoporosis in Red Angus cattle [75]. Similarly, in Texel sheep a mutation in the *SLC13A1* gene (1-bp deletion of T [g.25513delT] at the 107 bp position of exon 3) is directly associated with chondrodysplasia [76]. Thus, it appears that mutations in solute carrier families (e.g., in genes *SLC4A2* and *SLC13A1*), can cause bone disorders (e.g., chondrodysplasia) in livestock. In our research, we located the *SLC13A1* gene downstream from the most associated SNP (*chr4:88402194*) with width of hind leg. Therefore, we speculate that a mutation in gene *SLC13A1* may be responsible for bone disorders, especially in Simmental beef cattle. Hopefully, this insight can be used to help the cattle industry improve animal welfare and reduce economic losses.

We found that the gene *PLXNA4* (at 96 Mb) harbored 18 statistically significant SNPs. This gene belongs to the *plexin-A* subfamily of genes associated with the development of the nervous system. Moreover, the *PLXNA4* gene has the ability to regulate diverse processes, such as cell migration and neural growth cones (a domain with cytoskeletal instructions for actin) [77]. In humans, this gene plays a crucial role in denervation of skeletal muscles in motor neuron disease (MND) [78]. The function of this gene in cattle skeletal muscle is completely unknown; however, we hypothesize that the *PLXNA4* gene has a potential role in the development of skeletal muscle. However, more research is needed on functional relationships between the gene and skeletal development to test our hypothesis.

In the third sub-section we examined (at 97 Mb), we did not find any candidate genes from associations of SNPs with our measured traits. However, after exploring regions flanking SNPs, we identified two candidate genes (*CHCHD3* and *EXOC4*). We also identified the *EXOC4* gene in fourth sub-section of segment 1 we examined (at 98 Mb). Interestingly, all associated SNPs (n = 61) in sub-section 4 were located within this gene.

The *CHCHD3* gene is primarily expressed in mitochondria, especially in skeletal muscles [79]. Darshi et al. elucidated that the *CHCHD3* gene is essential for the formation of crista structure and in cellular metabolism [80], whereas Fan et al. suggested the potential relationship of this gene with back fat in pigs [81]. However, the function of the *CHCHD3* gene is almost unknown in cattle and so further research is needed to substantiate the association of this candidate gene with skeletal muscle and fat development.

The *EXOC4* gene which can cause schizophrenia in humans [82], is crucial for producing insulin and involved in protein and hormone metabolism. A few studies suggest an association between the *EXOC4* gene and meat quality in pigs and in Australian beef cattle. For example, Welzenbach et al. showed that the *EXOC4* gene may be a candidate gene associated with rate of drip loss in pigs [83]. However, more research is needed to identify the association between this candidate gene and meat quality in Simmental beef cattle.

In the sub-section 5 of Segment 1, we identified a very promising candidate gene (*MTPN*) related to skeletal muscle development, but this gene requires further validation. We also found a suite of statistically significant SNPs associated with the width of hind leg, providing candidate genes related to muscle development in the BTA4 chromosome. Our study highlights the possible role of the BTA4 region in molecular mechanism of muscle in Simmental cattle. Most of information we obtained from our research can be applied to breeding programs and thus help clarify mechanisms underlying the establishment of economically important traits in cattle.

## Supporting information

**S1 Table. Total reads of 44 samples (library) and their quality scores for next generation sequencing.**

(XLSX)

**S2 Table. Associated SNPs identified for width of hind leg trait.** The GWAS results of width of hind leg containing associated SNPs, p-values, and located gene(s).

(XLSX)

## Author Contributions

**Conceptualization:** Huijiang Gao, Junya Li.

**Data curation:** Zezhao Wang, Tianpeng Chang.

**Formal analysis:** Farhad Bordbar.

**Funding acquisition:** Lupei Zhang, Junya Li.

**Investigation:** Zezhao Wang, Lei Xu, Ling Xu.

**Methodology:** Farhad Bordbar.

**Project administration:** Farhad Bordbar, Huijiang Gao, Junya Li.

**Resources:** Bo Zhu, Lingyang Xu.

**Software:** Farhad Bordbar, Bo Zhu, Lingyang Xu.

**Supervision:** Just Jensen, Junya Li.

**Validation:** Farhad Bordbar, Just Jensen, Min Du.

**Visualization:** Farhad Bordbar.

**Writing – original draft:** Farhad Bordbar, Just Jensen.

**Writing – review & editing:** Farhad Bordbar, Just Jensen, Lingyang Xu.

## References

1. Seré C, Steinfeld H. World livestock production systems. FAO Animal Production and Health Paper 127. 1996; <http://www.fao.org/3/a-w0027e.pdf>
2. Harris DL. Breeding for Efficiency in Livestock Production: Defining the Economic Objectives. *J Anim Sci.* 1970; 30(6): 860–865.
3. Yuan Z, Li J, Li J, Gao X, Gao H, Xu S. Effects of DGAT1 gene on meat and carcass fatness quality in Chinese commercial cattle. *Mol Biol Rep.* 2012; 40(2): 1947–1954. <https://doi.org/10.1007/s11033-012-2251-2> PMID: 23143182
4. Albertí P, Panea B, Sañudo C, Olleta JL, Ripoll G, Ertbjerg P, et al. Live weight, body size and carcass characteristics of young bulls of fifteen European breeds. *Livestock Science.* 2008; 114: 19–30.
5. Swartz DR, Greaser ML, Cantino ME, Rhoads RP, Rathbone CR, Flann KL, et al. Applied Muscle Biology and Meat Science. In: Du M, McCormick RJ, editors. 1st ed. 2009. 360 p.
6. Tabor HK, Risch NJ, Myers RM. Candidate-gene approaches for studying complex genetic traits: practical considerations. *Nat Rev Genet.* 2002; 3: 391–397. <https://doi.org/10.1038/nrg796> PMID: 11988764
7. Zhu M, Zhao S. Candidate Gene Identification Approach: Progress and Challenges. *Int J Biol Sci.* 2007; 3(7): 420–427. <https://doi.org/10.7150/ijbs.3.420> PMID: 17998950
8. Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet.* 2005; 6(2): 95–108. <https://doi.org/10.1038/nrg1521> PMID: 15716906
9. Coon KD, Myers AJ, Craig DW, Webster JA, Pearson JV, Lince DH, et al. A high-density whole-genome association study reveals that APOE is the major susceptibility gene for sporadic late-onset Alzheimer's disease. *J Clin Psychiatry.* 2007; 68(4): 613–618. <https://doi.org/10.4088/jcp.v68n0419> PMID: 17474819
10. Gudmundsson J, Sulem P, Gudbjartsson DF, Blondal T, Gylfason A, Agnarsson BA, et al. Genome-wide association and replication studies identify four variants associated with prostate cancer susceptibility. *Nat Genet.* 2009; 41(10): 1122–1126. <https://doi.org/10.1038/ng.448> PMID: 19767754
11. Goddard ME, Hayes BJ. Mapping genes for complex traits in domestic animals and their use in breeding programmes. *Nat Rev Genet.* 2009; 10(6): 381–391. <https://doi.org/10.1038/nrg2575> PMID: 19448663
12. Maltecca C, Weigel KA, Khatib H, Cowan M, Bagnato A. Whole-genome scan for quantitative trait loci associated with birth weight, gestation length and passive immune transfer in a Holstein x Jersey cross-bred population. *Anim Genet.* 2009; 40: 27–34. <https://doi.org/10.1111/j.1365-2052.2008.01793.x> PMID: 19016677
13. McClure MC, Morsci NS, Schnabel RD, Kim JW, Yao P, Rolf MM, et al. A genome scan for quantitative trait loci influencing carcass, post-natal growth and reproductive traits in commercial Angus cattle. *Anim Genet.* 2010; 41: 597–607. <https://doi.org/10.1111/j.1365-2052.2010.02063.x> PMID: 20477797

14. Cole JB, Wiggans GR, Ma L, Sonstegard TS, Lawlor TJ, Crooker BA, et al. Genome-wide association analysis of thirty one production, health, reproduction and body conformation traits in contemporary U. S. Holstein cows. *BMC Genomics*. 2011; <https://doi.org/10.1186/1471-2164-12-408>
15. Visscher PM, Wray NR, Zhang Q, Sklar P, McCarthy MI, Brown MA, et al. 10 years of GWAS discovery: biology, function, and Translation. *Am J Hum Genet*. 2017; 101(1): 5–22. <https://doi.org/10.1016/j.ajhg.2017.06.005> PMID: 28686856
16. Reis-Filho JS. Next-generation sequencing. *Breast Cancer Res*. 2009; <https://doi.org/10.1186/bcr2431>
17. Sharma A, Cho Y, Choi BH, Chai HH, Park JE, Lim D. Limited representation of OMIA causative mutations for cattle in SNP databases. *Anim Genet*. 2017; 48: 369–370. <https://doi.org/10.1111/age.12534> PMID: 28116762
18. Shin DH, Lee HJ, Cho S, Kim HJ, Hwang JY, Lee CK, et al. Deleted copy number variation of Hanwoo and Holstein using next generation sequencing at the population level. *BMC Genomics*. 2014; <https://doi.org/10.1186/1471-2164-15-240>
19. Rubin CJ, Megens HJ, Barrio AM, Maqbool K, Sayyab S, Schwochow D, et al. Strong signatures of selection in the domestic pig genome. *Proc Natl Acad Sci U S A*. 2012; 109: 19529–19536.
20. Choi JW, Chung WH, Lee KT, Cho ES, Lee SW, Choi BH, et al. Whole-genome resequencing analyses of five pig breeds, including Korean wild and native, and three European origin breeds. *DNA Res*. 2015; 22: 259–267. <https://doi.org/10.1093/dnares/dsv011> PMID: 26117497
21. McClure MC, Morsci NS, Schnabel RD, Kim JW, Yao P, Rolf MM, et al. A genome scan for quantitative trait loci influencing carcass, post-natal growth and reproductive traits in commercial Angus cattle. *Anim Genet*. 2010; 41(6): 597–607. <https://doi.org/10.1111/j.1365-2052.2010.02063.x> PMID: 20477797
22. Nishimura S, Watanabe T, Mizoshita K, Tatsuda K, Fujita T, Watanabe N, et al. Genome-wide association study identified three major QTL for carcass weight including the PLAG1-CHCHD7 QTN for stature in Japanese Black cattle. *BMC Genet*. 2012; <https://doi.org/10.1186/1471-2156-13-40>
23. Setoguchi K, Furuta M, Hirano T, Nagao T, Watanabe T, Sugimoto Y, et al. Cross-breed comparisons identified a critical 591-kb region for bovine carcass weight QTL (CW-2) on chromosome 6 and the Ile-442-Met substitution in NCAPG as a positional candidate. *BMC Genet*. 2009; <https://doi.org/10.1186/1471-2156-10-43>
24. Takasuga A, Watanabe T, Mizoguchi Y, Hirano T, Ihara N, Takano A, et al. Identification of bovine QTL for growth and carcass traits in Japanese Black cattle by replication and identical-by-descent mapping. *Mamm Genome*. 2017; 18(2): 125–136.
25. Lindholm-Perry AK, Sexten AK, Kuehn LA, Smith TP, King DA, Shackelford SD, et al. Association, effects and validation of polymorphisms within the NCAPG—LCORL locus located on BTA 6 with feed intake, gain, meat and carcass traits in beef cattle. *BMC Genet*. 2011; <https://doi.org/10.1186/1471-2156-12-103>
26. Riggs PK, Owens KE, Rexroad CE, AmaraL MEJ, Womack JE. Development and initial characterization of a *Bos taurus* x *B. Gaurus* interspecific hybrid backcross panel. *J Hered*. 1997; 88: 373–379. <https://doi.org/10.1093/oxfordjournals.jhered.a023121> PMID: 9378912
27. Threadgill DS, Threadgill DW, Moll YD, Weiss JA, Zhang N, Davey HW, et al. Syntenic assignment of human chromosome 1 homologous loci in the bovine. *Genomics*. 1994; 22: 626–630. <https://doi.org/10.1006/geno.1994.1436> PMID: 8001974
28. Bishop MD, Kappes SM, Keele JW, Stone RT, Sunden SL, Hawkins GA, et al. A genetic linkage map for cattle. *Genetics*. 1994; 36: 619–639.
29. Charlier C, Coppieters W, Farnir F, Grobet L, Leroy PL, Michaux C, et al. The MH gene causing double-muscling in cattle maps to bovine chromosome 2. *Mamm Genome*. 1995; 6(11): 788–792. <https://doi.org/10.1007/bf00539005> PMID: 8597635
30. Beaver JE, Fisher SR, Guérin G, Lewin HA. Mapping of eight human chromosome 1 orthologs to cattle chromosomes 3 and 16. *Mamm Genome*. 1997; 8(7): 533–536. <https://doi.org/10.1007/s003359900493> PMID: 9196004
31. Smith TP, Casas E, Rexroad CE 3rd, Kappes SM, Keele JW. Bovine CAPN1 maps to a region of BTA29 containing a quantitative trait locus for meat tenderness. *J Anim Sci*. 2000; 78(10): 2589–2594. <https://doi.org/10.2527/2000.78102589x> PMID: 11048924
32. Cheong HS, Yoon DH, Park BL, Kim LH, Bae JS, Namgoong S, et al. A single nucleotide polymorphism in CAPN1 associated with marbling score in Korean cattle. *BMC Genet*. 2008; <https://doi.org/10.1186/1471-2156-9-33>
33. Piedrafitra J, Quintanilla R, Sanudo C, Olleta JL, Campo MM, Panea B, et al. Carcass quality of 10 beef cattle breeds of the Southwest of Europe in their typical production systems. *Livest Prod Sci*. 2003; 82(1): 1–13.

34. Booth FW, Kelso JR. Effect of hind-limb immobilization on contractile and histochemical properties of skeletal muscle. *Pflugers Archiv*. 1973; 342(3): 231–238. <https://doi.org/10.1007/bf00591371> PMID: 4270552
35. Oddy VH, Herd RM, McDonagh MB, Woodgate R, Quinn CA, Zirkler K. Effect of divergent selection for yearling growth rate on protein metabolism in hind-limb muscle and whole body of Angus cattle. *Livest Prod Sci*. 1998; 56(3): 225–231.
36. Allen DL, Linderman JK, Roy RR, Bigbee AJ, Grindeland RE, Mukku V, et al. Apoptosis: a mechanism contributing to remodeling of skeletal muscle in response to hindlimb unweighting. *Am J Physiol*. 1997; <https://doi.org/10.1152/ajpcell.1997.273.2.C579>
37. Boisclair YR, Bauman DE, Bell AW, Dunshea FR, Harkins M. Nutrient utilization and protein turnover in the hindlimb of cattle treated with bovine somatotropin. *J Nutr*. 1994; 124(5): 664–673. <https://doi.org/10.1093/jn/124.5.664> PMID: 8169658
38. Dutra F, Carlsten J, Ekman S. Hind limb skeletal lesions in 12-month-old bulls of beef breeds. *Zentralbl Veterinarmed A*. 1999; 46(8): 489–508. <https://doi.org/10.1046/j.1439-0442.1999.00241.x> PMID: 10596288
39. McDonagh MB, Fernandez C, Oddy VH. Hind-limb protein metabolism and calpain system activity influence post-mortem change in meat quality in lamb. *Meat Sci*. 1999; 52(1): 9–18. [https://doi.org/10.1016/s0309-1740\(98\)00143-0](https://doi.org/10.1016/s0309-1740(98)00143-0) PMID: 22062138
40. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007; 81(3): 559–575. <https://doi.org/10.1086/519795> PMID: 17701901
41. Fujimoto A, Nakagawa H, Hosono N, Nakano K, Abe T, Boroevich KA, et al. Whole genome sequencing and comprehensive variant analysis of a Japanese individual using massively parallel sequencing. *Nature Genet*. 2010; 42: 931–936. <https://doi.org/10.1038/ng.691> PMID: 20972442
42. Browning BL, Browning SR. Genotype imputation with millions of reference samples. *Am J Hum Genet*. 2016; 98: 116–126. <https://doi.org/10.1016/j.ajhg.2015.11.020> PMID: 26748515
43. Fuchsberger C, Abecasis GR, Hinds DA. Minimac2: faster genotype imputation. *Bioinformatics*. 2014; 31: 782–784. <https://doi.org/10.1093/bioinformatics/btu704> PMID: 25338720
44. Aulchenko YS, Ripke S, Isaacs A, van Duijn CM. GenABEL: An R library for genome-wide association analysis. *Bioinformatics*. 2007; 23(10): 1294–1296. <https://doi.org/10.1093/bioinformatics/btm108> PMID: 17384015
45. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*. 2005; 21: 263–265. <https://doi.org/10.1093/bioinformatics/bth457> PMID: 15297300
46. Li F, Chen B, Xu K, Wu J, Song W, Bancroft I, et al. Genome-wide association study dissects the genetic architecture of seed weight and seed quality in rapeseed (*Brassica napus L.*). *DNA Res*. 2014; 21(4): 355–367. <https://doi.org/10.1093/dnares/dsu002> PMID: 24510440
47. Bolormaa S, Hayes BJ, Savin K, Hawken R, Barendse W, Arthur PF, et al. Genome-wide association studies for feedlot and growth traits in cattle. *J Anim Sci*. 2011; 89(6): 1684–1697. <https://doi.org/10.2527/jas.2010-3079> PMID: 21239664
48. Bolormaa S, Pryce JE, Reverter A, Zhang Y, Barendse W, Kemper K, et al. A multi-trait, meta-analysis for detecting pleiotropic polymorphisms for stature, fatness and reproduction in beef cattle. *PLoS Genet*. 2014; <https://doi.org/10.1371/journal.pgen.1004198>
49. Wu X, Fang M, Liu L, Wang S, Liu J, Ding X, et al. Genome wide association studies for body conformation traits in the Chinese Holstein cattle population. *BMC Genomics*. 2013; <https://doi.org/10.1186/1471-2164-14-897>
50. Abo-Ismael MK, Vander Voort G, Squires JJ, Swanson KC, Mandell IB, Liao X, et al. Single nucleotide polymorphisms for feed efficiency and performance in crossbred beef cattle. *BMC Genet*. 2014; <https://doi.org/10.1186/1471-2156-15-14>
51. Feugang JM, Kaya A, Page GP, Chen L, Mehta T, Hirani K, et al. Two-stage genome-wide association study identifies integrin beta 5 as having potential role in bull fertility. *BMC Genomics*. 2009; <https://doi.org/10.1186/1471-2164-10-176>
52. Duijvesteijn N, Knol EF, Merks JW, Crooijmans RP, Groenen MA, Bovenhuis H, et al. A genome-wide association study on androstenedione levels in pigs reveals a cluster of candidate genes on chromosome 6. *BMC Genet*. 2010; <https://doi.org/10.1186/1471-2156-11-42>
53. Barendse W, Reverter A, Bunch RJ, Harrison BE, Barris W, Thomas MB. A validated whole-genome association study of efficient food conversion in cattle. *Genetics*. 2007; 176(3): 1893–1905. <https://doi.org/10.1534/genetics.107.072637> PMID: 17507676
54. Kolbehdari D, Wang Z, Grant JR, Murdoch B, Prasad A, Xiu Z, et al. A whole genome scan to map QTL for milk production traits and somatic cell score in Canadian Holstein bulls. *J Ani Breed Genet*. 2009; 126(3): 216–227.

55. Settles M, Zanella R, McKay SD, Schnabel RD, Taylor JF, Whitlock R, et al. A Whole Genome Association analysis identifies loci associated with *Mycobacterium avium* subsp. *paratuberculosis* infection status in US Holstein Cattle. *Anim Genet*. 2009; 40(5): 655–662. <https://doi.org/10.1111/j.1365-2052.2009.01896.x> PMID: 19422364
56. Liu Y, Li J, Zhang F, Qin W, Yao G, He X, et al. Molecular cloning and characterization of the human ASB-8 gene encoding a novel member of ankyrin repeat and SOCS box containing protein family. *Biochem Biophys Res Commun*. 2003; 300(4): 972–979. [https://doi.org/10.1016/s0006-291x\(02\)02971-6](https://doi.org/10.1016/s0006-291x(02)02971-6) PMID: 12559969
57. Hancock DL, Wagner J, and Anderson D. Effects of estrogens and androgens on animal growth. In: Pearson A, Dutton T, editors. *Advances in Meat Research*. vol 7. Essex, UK: Elsevier Science; 1991. p. 255. <http://agris.fao.org/agris-search/search.do?recordID=US9442858>
58. Kohroki J, Fujita S, Itoh N, Yamada Y, Imai H, Yumoto N, et al. ATRA-regulated ASB-2 gene induced in differentiation of HL-60 leukemia cells. *FEBS Lett*. 2001; 505(2): 223–228. [https://doi.org/10.1016/s0014-5793\(01\)02829-0](https://doi.org/10.1016/s0014-5793(01)02829-0) PMID: 11566180
59. McDanel TG, Hancock DL, Moody DE. Altered mRNA abundance of ASB15 and four other genes in skeletal muscle following administration of  $\beta$ -adrenergic receptor agonists. *Physiol Genomics*. 2004; 16(2): 275–283. <https://doi.org/10.1152/physiolgenomics.00127.2003> PMID: 14645738
60. McDanel TG, Hannon K, Moody DE. Ankyrin repeat and SOCS box protein 15 regulates protein synthesis in skeletal muscle. *Am J Physiol Regul Integr Comp Physiol*. 2006; 290(6): R1672–R1682. <https://doi.org/10.1152/ajpregu.00239.2005> PMID: 16424087
61. McDanel TG, Spurlock DM. Ankyrin repeat and suppressor of cytokine signaling (SOCS) box-containing protein (ASB) 15 alters differentiation of mouse C2C12 myoblasts and phosphorylation of mitogen-activated protein kinase and Akt1. *J Anim Sci*. 2008; 86(11): 2897–2902. <https://doi.org/10.2527/jas.2008-1076> PMID: 18641171
62. Ackrell BA, Maguire JJ, Dallman PR, Kearney EB. Effect of iron deficiency on succinate- and NADH-ubiquinone oxidoreductases in skeletal muscle mitochondria. *J Biol Chem*. 1984; 259: 10053–10059. PMID: 6432778
63. Marui T, Funatogawa I, Koishi S, Yamamoto K, Matsumoto H, Hashimoto O, et al. The NADH-ubiquinone oxidoreductase 1 alpha subcomplex 5 (NDUFA5) gene variants are associated with autism. *Acta Psychiatr Scand*. 2011; 123(2): 118–124. <https://doi.org/10.1111/j.1600-0447.2010.01600.x> PMID: 20825370
64. Seong J, Yoon H, Kong HS. Identification of microRNA and target gene associated with marbling score in Korean cattle (Hanwoo). *Genes & Genomics*. 2016; 38(6): 529–538.
65. Keogh K, Kenny DA, Kelly AK, Cormican P, Waters S. Feed restriction and subsequent re-alimentation affect the expression of oxidative phosphorylation genes in skeletal muscle of Holstein Friesian bulls. 10th World Congress on Genetics Applied to Livestock Production. 2014. <http://www.wcgalp.org/proceedings/2014/feed-restriction-and-subsequent-re-alimentation-affect-expression-oxidative>
66. Cenik BK, Garg A, McAnally JR, Shelton JM, Richardson JA, Bassel-Duby R, et al. Severe myopathy in mice lacking the MEF2/SRF-dependent gene *leiomod*-3. *J Clin Invest*. 2015; 125(4): 1569–1578. <https://doi.org/10.1172/JCI80115> PMID: 25774500
67. Chereau D, Boczkowska M, Skwarek-Maruszewska A, Fujiwara I, Hayes DB, Rebowksi G, et al. *Leiomod* is an actin filament nucleator in muscle cells. *Science*. 2008; 320(5873): 239–243. <https://doi.org/10.1126/science.1155313> PMID: 18403713
68. Nworu CU, Kraft R, Schnurr DC, Gregorio CC, Krieg PA. *Leiomod* 3 and *tropomod* 4 have overlapping functions during skeletal myofibrillogenesis. *J Cell Sci*. 2015; 128(2): 239–250. <https://doi.org/10.1242/jcs.152702> PMID: 25431137
69. Pappas CT, Mayfield RM, Henderson C, Jamilpour N, Cover C, Hernandez Z, et al. Knockout of *LMOD2* results in shorter thin filaments followed by dilated cardiomyopathy and juvenile lethality. *Proc Natl Acad Sci U S A*. 2015; 112(44): 13573–13578. <https://doi.org/10.1073/pnas.1508273112> PMID: 26487682
70. Yuen M, Sandaradura SA, Dowling JJ, Kostyukova AS, Moroz N, Quinlan KG, et al. *Leiomod*-3 dysfunction results in thin filament disorganization and nemaline myopathy. *J Clin Invest*. 2014; 124(11): 4693–4708. <https://doi.org/10.1172/JCI75199> PMID: 25250574
71. Yu SL, Chung HJ, Sang BC, Park CS, Lee JH, Yoon DH, et al. Identification of differentially expressed genes in distinct skeletal muscles in cattle using cDNA microarray. *Anim Biotechnol*. 2007; 18(4): 275–285. <https://doi.org/10.1080/10495390701413391> PMID: 17934901
72. Li K, Ma YB, Zhang Z, Tian YH, Xu XL, He YQ, et al. Upregulated IQUB promotes cell proliferation and migration via activating Akt/GSK3 $\beta$ / $\beta$ -catenin signaling pathway in breast cancer. *Cancer Med*. 2018; 7(8): 3875–3888. <https://doi.org/10.1002/cam4.1568> PMID: 29968965



73. Pinto CS, Khandekar A, Bhavana R, Kiesel P, Pigino G, Sonawane M. Microridges are apical projections formed of branched F-actin networks that organize the glycan layer. *Biorxiv*. 2018; <https://doi.org/10.1101/442871>
74. Hästbacka J, Superti-Furga A, Wilcox WR, Rimoin DL, Cohn DH, Lander ES. Atelosteogenesis type II is caused by mutations in the diastrophic dysplasia sulfate-transporter gene (DTDST): evidence for a phenotypic series involving three chondrodysplasias. *Am J Hum Genet*. 1996; 58(2): 255–262. PMID: 8571951
75. Meyers SN, McDanel TG, Swist SL, Marron BM, Steffen DJ, O'Toole D, et al. A deletion mutation in bovine SLC4A2 is associated with osteopetrosis in Red Angus cattle. *BMC Genomics*. 2010; <https://doi.org/10.1186/1471-2164-11-337>
76. Zhao X, Onteru SK, Piripi S, Thompson KG, Blair HT, Garrick DJ, et al. In a shake of a lamb's tail: using genomics to unravel a cause of chondrodysplasia in Texel sheep. *Anim Genet*. 2012; 1:9–18.
77. Gutekunst CA, Stewart EN, Franz CK, English AW, Gross RE. PlexinA4 Distribution in the Adult Rat Spinal Cord and Dorsal Root Ganglia. *J Chem Neuroanat*. 2012; 44(1): 1–13. <https://doi.org/10.1016/j.jchemneu.2012.03.002> PMID: 22465808
78. De Winter F, Vo T, Stam FJ, Wisman LA, Bär PR, Niclou SP, et al. The expression of the chemorepellent Semaphorin 3A is selectively induced in terminal Schwann cells of a subset of neuromuscular synapses that display limited anatomical plasticity and enhanced vulnerability in motor neuron disease. *Mol Cell Neurosci*. 2006; 32(1–2): 102–117. <https://doi.org/10.1016/j.mcn.2006.03.002> PMID: 16677822
79. Hwang H, Bowen BP, Lefort N, Flynn CR, De Filippis EA, Roberts C. Proteomics analysis of human skeletal muscle reveals novel abnormalities in obesity and type 2 diabetes. *ADA Journal*. 2010; 59(1): 33–42.
80. Darshi M, Mendiola VL, Mackey MR, Murphy AN, Koller A, Perkins GA, et al. ChChd3, an inner mitochondrial membrane protein, is essential for maintaining crista integrity and mitochondrial function. *J Biol Chem*. 2011; 286(4): 2918–2932. <https://doi.org/10.1074/jbc.M110.171975> PMID: 21081504
81. Fan B, Onteru SK, Du ZQ, Garrick DJ, Stalder KJ, Rothschild MF. Genome-wide association study identifies loci for body composition and structural soundness traits in pigs. *PLoS One*. 2011; <https://doi.org/10.1371/journal.pone.0014726>
82. Lee SA, Chan CH, Chen TC, Yang CY, Huang KC, Tsai CH, et al. Protein interactome with sub-network analysis and hub prioritization. *BMC Bioinformatics*. 2009; <https://doi.org/10.1186/1471-2105-10-114>
83. Welzenbach J, Neuhoff C, Heidt H, Cinar MU, Looft C, Schellander K, et al. Integrative analysis of metabolomics, proteomic and genomic data to reveal functional pathways and candidate genes for drip loss in pigs. *Int J Mol Sci*. 2016; <https://doi.org/10.3390/ijms17091426>