Multiple association analysis strategies identified loci and candidate genes for body size on three growth stages in Simmental beef cattle

Bingxing An¹#, Lei Xu¹#, Jiangwei Xia², Tianpeng Chang¹, Xiaqiao Wang¹, Lingyang Xu¹, Lupei Zhang¹, Xue Gao¹, Junya Li¹, and Huijiang Gao¹*

Authors' affiliations:
¹Institute of Animal Science, Chinese Academy of Agricultural Science, Beijing, 100193, China.
²Institute of Basic Medical Sciences, Westlake Institute for Advanced Study, Hangzhou, 310000, China.

#These authors contributed equally to this work.

*Corresponding authors:

Dr. Huijian Gao
E-mail: gaohuijiang@caas.cn

Emails of the other contributor:

Bingxing An: ABX2HF@126.com
Lei Xu: xuleirock@163.com
Jiangwei Xia: jiangwei_xia06@126.com
Tianpeng Chang: 13051378730@163.com
Xiaqiao Wang: longmao7@live.com
Lingyang Xu: xulingyang@163.com
Lupei Zhang: lpzhang@iascaas.net.cn
Xue Gao: gaoxue76@126.com
Junya Li: jll@iascaas.net.cn
Abstract

Background: Body size traits as one of the main breeding selection criteria have long since being widely used to monitor cattle growth and evaluate the selection response. Here the volume of body size is indicated by body height (BH), body length (BL), hip height (HH), heart size (HS), abdominal size (AS) and cannon bone size (CS). We performed genome-wide association studies (GWAS) for these traits to a broad spectrum of three growth stages (months 6, 12 and 18 after birth) under three statistical models: single-trait GWAS, multi-trait GWAS and LONG-GWAS. The whole genomic single nucleotide polymorphisms (SNPs) were obtained from the Illumina Bovine HD 770K BeadChip genotype on 1217 individuals.

Results: In total, 19, 29, and 10 significant SNPs were identified by the three models, respectively. While 21 genes among in these loci appeared to be promising candidate genes, including SOX2, SNRPD1, RASGEF1B, EFNA5, PTBP1, SNX9, SV2C, PKDCC, SYNDIG1, AKR1E2 and PRIM2 detected by single-trait analyze; SLC37A1, LAP3, PCDH7, MANEA and LHCG detected by multi-trait analyze; P2RY1, MPZL1, LINGO2, CMIP and WSCD1 detected by LONG-GWAS.

Conclusions: Multiple association analysis strategies were performed for six growth traits on each stage. This research could offer valuable insights to further explore the potential mechanism of growth traits in Simmental beef cattle.

Keywords: Simmental Beef cattle, Genome-wide association studies, body size, candidate genes, Bovine HD 770K SNP
Background

Beef cattle production plays an increasingly important role in Chinese agribusiness and Simmental breed accounts for more than 70% of beef-producing herds. The fact that body size was frequently selected by beef producers to monitor growth and track the developmental progress of each animal throughout the fattening period in farming [1, 2]. While these economic traits are of interest to health [3], longevity [4] and increasing profit in efficiency of feed and management [5–7], which have received growing attention for more than a century. Besides in human, additive genetic effect explains 81% of the variation in height [8], and the heritability for both HH and HS in cattle is 0.33-0.4 [9]. Bouwman et al. reported that the lead variants in significant regions explained at most 13.8% of the phenotypic variance in their meta-analysis for cattle size using 58,265 cattle from 17 populations [10]. In addition, body linear measurements, specifically body height (BH) and hip height (HH), highly reliable and accurate indicators for body weight, has been explored to be much easier to measure than body weight in daily production [11]. The depth of height size (HS) is an indication of good feed conversion and carcass information [12]. Dairy cows with higher HH will subsequently have better milk performance [13]. However, little knows about the molecular mechanisms of body size traits in Chinese Simmental beef cattle.

Genome-wide association studies (GWAS) as a robust statistic tool are being broadly asked to identify candidate genes with significant SNPs in production traits [14–16], growth traits [17, 18], carcass quality traits and fertility traits [19, 20]. By focusing on beef cattle, numerous SNPs, genes and haplotype blocks were discovered associated with growth, but the current GWAS-based studies mainly focus on only one growth stage [21], such as weaning size [22], yearling weight and stature in slaughter [23]. Besides in different growth stages of beef cattle, loci controlling growth traits may be unlike, and some of which may constantly control the traits during the whole life process, so it is more reasonable to perform GWAS for the growth traits on each stage separately. While multi-trait methods have been developed to increase statistical power and identify pleiotropic loci in GWAS [24]. Furthermore, the longitudinal GWAS model jointly considering all time points when performing association has been made to assess whether some significant SNPs are associated with the process that a trait develops over time [25], and this method showed powerful for identifying these time-dependent and consistent loci [26]. The multi-trait GWAS and LONG-GWAS not be a
replacement for the single-trait GWAS, but rather as suggestive complementary to it. While dissecting the genetic mechanism of inter-individual variation in body size with these methods might provide new sights that can help manipulate cattle growth and production.

In this research, six body size traits from entering farm to slaughter have become routine measurements during the past 10 years, which provides valuable resources to study the complete growing period. The purpose of our study was tantamount to expecting a comprehensive analysis of candidate genes and QTL regions associated with growth traits through the entire growth and development process by conducting three GWAS approaches in Simmental beef cattle. Our results could offer valuable insights to further explore the potential mechanism of growth traits in Simmental beef cattle.

**Results**

**Population stratification assessment**

Figure 1 shows that the population stratification of the Simmental population based on the PCA was divided into five separate clusters, demonstrating an obvious stratification in the reference population. Population stratification caused by different genetic influences and breeding conditions, as a potential confounder, was corrected through a significance testing. The Manhattan plots and “Q-Q” plots are shown in Figure 2 and Figure 3, while “Q-Q” plots with majority of points revolving around the diagonal line show that there is no inflation or systematic bias in this research, which caused by the GWAS model considered the population structure in full and a few SNPs were associated with the target traits.

**Summary of significant loci identified by three approaches**

Briefly, we found 45, 66 and 19 SNPs significantly were associated with six body size traits by single-trait GWAS, multi-trait GWAS and LONG-GWAS. There were no significant loci for single-trait BH6 AS6, CS6 and LONG-AS. The corresponding Manhattan plots are shown in Figure 2. In addition, ten SNPs were associated with at least one of six traits and eight SNPs show strong association with these traits in at least one of three models. While according to their biological function, 21 suggestive genes were selected as candidate genes and some details of them, including their positions in the genome, the nearest reported genes, the MAF and the p values were listed in Table 2.
SNPs identified by single-trait GWAS

A total of 45 SNPs achieved genome-wide significance associated with at least one of the six traits, with the p-value ranging from 9.99×10^{-6} (BovineHD0700018941 for BL18) to 2.11×10^{-8} (BovineHD0200014365 for HS18), and the MAF ranging from 0.003 (BovineHD0700034055) to 0.497 (BovineHD2600012755). Among them, two SNPs near SOX2 on BTA1, was related to gene expression in the liver and stomach [37]. On BTA2, two loci in the 0.69 Kb region were significantly associated with single-trait HS18 and one of them (BovineHD0200014365) was also associated with multi-trait18. On BTA6, three SNPs in the 0.46 Mb region were located near RASGEF1B, a guanine nucleotide exchange factor with specificity for Rap2A [38]. Besides on BTA7, one SNP (BovineHD0700034055) was associated with single-trait HS12 and single-trait AS12, namely EFNA5 and another SNP (BovineHD0700012966) was associated with single-trait HH6 and multi-trait6, namely PTBP1. On BTA10, one SNP (BovineHD1000002378) was associated with single-trait BH6 and multi-trait6, namely SV2C. While on BTA11, four loci in 0.04 Mb region were associated with the single-trait HS12 and one of them (BovineHD1100007368) also showed association with multi-trait12, all of which were near PKDCC. On BTA13, SNP BovineHD1300012489 and BovineHD1300012894 were associated with single-trait BH18 and single-trait AS18, respectively. Besides them also showed strong association with multi-trait18. On BTA23, two SNPs were located within PRIM2, which encodes the DNA primase, an enzyme that plays a key role in the replication of DNA [39].

SNPs identified by multi-trait GWAS

The multi-trait GWAS identified 66 SNPs within or near 36 genes that were distributed on 21 chromosomes, including 8 loci that also were identified in mentioned single-trait GWAS, which indicated that these loci suggestively regulate the development of the body growth (Table 2). Among them, two promising loci in the 11.4 Kb region were detected, namely SLC37A1. On BTA6, two suggestive loci were detected, one near LAP3 that associated with multi-trait12 and another near PCDH7. Two genome-wide loci were identified at 0.02 Mb region on BTA9, namely MANEA. Besides four promising loci within LHCGR were detected at 0.04 Mb region on BTA11. Moreover, thirteen loci near AKR1E2 at 0.03Mb region were identified on BTA13, which also were detected in mentioned single-trait analysis.
SNPs identified by LONG-GWAS

Nineteen loci were identified in the LONG-GWAS, including three significant loci on 12 chromosomes (Table 2). Among them, two suggestive loci near P2RY1 in the 8.3Kb region were detected, while the later (BovineHD0100032742) also was associated with LONG-CS, LONG-BH and LONG-HH. Another two loci near MPZL1 in the 4.2Kb region were identified and the latter SNP (BTA-68271-no-rs) also showed association with LONG-HH and LONG-BH. Besides, one suggestive SNP near LINGO2 on BTA8 was associated with LONG-HS, which encoded a transmembrane protein belonging to the LINGO/LERN protein family which contains four homologs in the human genome (LINGO1-4) [40]. Additionally, four promising loci in the 0.03 Kb region were detected to associate with LONG-HH, namely CMIP, a key gene in T-cell signaling pathway. On BTA19, a suggestive locus near WSCD1 was associated with LONG-CS. No loci were associated with AS in our research.

Discussion

We performed single-trait GWAS, multi-trait GWAS, and LONG-GWAS for six body size traits on three growth stages in Simmental beef cattle. However, the three methods yielded different results with few shared loci, the reasons for this condition are discussed below: first, which is likely caused by the restricted dataset in single-trait GWAS and LONG-GWAS analysis. One universal phenomenon that cannot be ignored is that growth traits are controlled by polygenes with small effects. Then each method has its advantage to identify distinct loci. For example, single-trait GWAS is robust to detect trait-specific QTL and multi-trait GWAS is efficient for mapping pleiotropic QTL [41], while LONG-GWAS can improve the detection power for time-dependent and consistent loci [42]. Thus, combining these three GWAS methods was expected to analyze the genetic mechanism of the body traits of beef cattle more comprehensively and convincingly. In addition, since many complex traits have a similar architecture across diverse species [7], we tried to compare some of our significant genes with the previous reports about the same genes and their association with growth. As a result, 21 suggestive genes were considered as candidate genes, which involved in development or associated with growth in cattle, swine, human and mice studies.

Candidate Genes Identified by Single-trait GWAS

On BTA1, two SNPs near SOX2 (SRY-Box 2), was reported previously encoded a transcription factor
involved in the regulation of embryonic development [43, 44]. While an important paralog of this gene is SOX17, which has a positive effect on the growth traits of cattle, and the conserved regions of this gene in human genome is closely related to body development [45]. On BTA2, two SNPs were near SNRPD1 (Small Nuclear Ribonucleoprotein D1 Polypeptide), which is a member of the ghrelin receptor family, and the encoded protein is involved in zinc-dependent signaling in epithelial tissue [46]. Besides on BTA6, variations near RASGEF1B (RasGEF Domain Family Member 1B) have been associated with body height [47], as well as body height is positively correlated with calcium absorption efficiency that is important determinant of calcium balance [48]. On BTA7, a SNP near EFNA5 (Ephrin A5) was identified to associate with two traits (HS and AS) in the same stage, which has been reported as candidate gene for growth traits in broiler chicken cross [49]. Another one near PTBP1 (Polypyrimidine Tract Binding Protein 1) was identified showing genome-wide association with growth traits at 6 month old by both single-trait and multi-trait GWAS, whose expression level determined the release of insulin, thereby affected body development [50]. On BTA9, a SNP (BovineHD0900027283) located in SNX9 (Sorting Nexin 9), as an olfactory receptor, which was previously identified to be associated with growth traits in Yorkshire pig [51]. While the SNP near SV2C (Synaptic Vesicle Glycoprotein 2C) showed association with BH by both single-trait and multi-trait GWAS, which was reported to modulate dopamine release in neural and endocrine cells [52]. On BTA11, PKDCC (Protein Kinase Domain Containing, Cytoplasmic) was associated with HS in both single-trait and multi-trait analysis, which was reported as a region associated with bone density in human [53]. On BTA13, SYNDIG1 (Synapse Differentiation Inducing 1) has been reported as a factor influencing the final weight and backfat thickness of Landrace pigs [54]. While the variants of AKR1E2 (Aldo-Keto Reductase Family 1 Member E2) were previously associated with body length and girth in cattle [55]. On BTA23, the PRIM2 (DNA Primase Subunit 2) was reported previously having association with body weight changes and measurement traits in pig population [56, 57].

**Single-trait GWAS versus multi-trait GWAS**

Multiple-trait analysis of linkage experiments has been reported to significantly enhance the power to detect common SNPs across traits [58, 59]. Therefore, the multi-trait analysis was a complementary part to single-trait GWAS rather than replacement for it. In the single-trait GWAS, the minimum p values for three stages were 3.90E-07, 9.92E-07 and 2.11E-08 respectively; whereas these three values decreased to 8.23E-13, 1.73E-09 and 3.84E-10 in the multi-trait GWAS, respectively. While multi-
trait GWAS also identified some critical loci as follows. On BTA1, the *SLC37A1* (solute carrier family 37, member A1) gene, which encodes a glucose-6-phosphate transporter that is involved in the homeostasis of blood glucose [60], was previously proposed as the best candidate mutations for milk production traits [61]. On BTA6, *LAP3* (Leucine Aminopeptidase 3) was reported to play vital roles in the regulation of hormone levels and protein maturation. While some research indicated putative regulatory elements in the *PCDH7* (Protocadherin 7) gene that could have a role in RFI (residual feed intake) variation in Nelore cattle [62]. In addition, *MANEA* (Mannosidase Endo-Alpha), acting in proteolysis, has been shown association with birth weight in Canchim beef cattle [17]. On BTA11, a mutation in the *LHCGR* (Luteinizing Hormone/Chorionic Gonadotropin Receptor) gene was as the cause of EFS (empty follicle syndrome) [63].

**Single-trait GWAS versus LONG-GWAS**

We used LONG-GWAS that utilized multiple phenotype measurements for each individual [35]. While one disadvantage of which is that the significant signal may be overwhelmed by putting all time point’s data together if QTL effects vary during time stage [64]. In this study, these time-specific expressed QTL identified by the singer-marker GWAS were not detected by LONG-GWAS. However LONG-GWAS also detected some significant functional loci as follows. On BTA1, *P2RY1* (Purinergic Receptor P2Y1), a candidate gene for affecting the level of serum Ca^{2+}, encoded for a member of the family of G protein-coupled receptors that works as receptor for extracellular ATP and ADP [65]. On BTA3, *MPZL1* (Myelin Protein Zero Like 1) could significantly enhance the migratory and metastatic potential of the HCC (hepatocellular carcinoma) cells through the mechanisms including phosphorylation and activation of the pro-metastatic protein [66]. Besides on BTA8, *LINGO2* (Leucine Rich Repeat And Ig Domain Containing 2), with expression in the central nervous system in mouse embryos, has been observed to show association with body mass in a cohort of elderly Swedes [67]. On BTA18, *CMIP* (C-Maf Inducing Protein), a candidate gene for reading-related traits, was also associated with plasma lipoprotein levels [68]. Moreover, *WSCD1* (WSC Domain Containing 1), encoded a protein with sulfotransferase activity and involved the glucose metabolism, was a candidate gene for feed efficiency and feeding behaviors in a White Duroc × Erhualian F2 population [69].

**Conclusions**

In conclusion, a total of 58 SNPs corresponding to 21 genes were identified to be associated with six
body size traits at 6, 12 and 18 months. Future studies characterizing the functions of these candidate genes could provide valuable knowledge to uncover the genetic architecture underlying growth traits in Simmental beef cattle.

Methods

Resource Population and Phenotypes Collection

Simmental beef cattle (more than 2000 animals born in 2008-2015) were established in Ulgai, Xilingole League, Inner Mongolia of China. Six body size traits at three growth stages (6, 12, 18 months after birth) was measured simultaneously for each individual and blood sample were collected when the regular quarantine inspection of the farms was conducted, while the collection procedures were conducted in strict compliance with the guidelines established by the Ministry of Agriculture of China. Some descriptive statistics and heritability estimate of six traits at three growth stages are presented in Table 1.

Genotyping and quality control

Genomic DNA was isolated from blood samples using the TIANamp Blood DNA Kit (Tiangen Biotech Co.Ltd.,Beijing, China). DNA quality was considered eligible when the A260/A280 ratio ranged from 1.8 to 2.0. Genotyping was performed with the Illumina BovineHD Beadchip (Illumina Inc., San Diego, CA, USA) and the PLINK v1.07 software was used to conduct quality control [27]. In this study, animals with a call rate (<0.9) were discarded. SNPs were selected the following standards, including minor allele frequency (<0.01), genotype appearances (<0.05) and Hardy-Weinberg equilibrium values (p<1×10⁻⁶). Finally, 1217 Simmental cattle were generated for the analysis, which consisted of 671192 SNPs with an average distance of 3 kb.

Single-trait GWAS

We used the compressed mixed linear model (CMLM) for single-trait GWAS because it reduced computing time by clustering individuals into groups, increased power in QTN detection through eliminating the need to re-compute variance components and effectiveness in correcting the inflation from polygenic background and controlling the bias of population stratification [28, 29]. Briefly, a principal components analysis (PCA) was performed and a kinship matrix was calculated using the Genome Association and Prediction Integrated Tool (GAPIT) package in R v3.4.2 [30]. The Q matrix
was reflected by the PCA to revise the effects of population structure and the $K$ matrix was calculated by the VanRaden algorithm to replace the incomplete genealogy [31]. This model is as follows:

$$y = W\nu + X\beta + Zu + e$$

where $y$ is a vector of the observed phenotypes; Variable $W$ was coded as 0, 1 and 2 corresponding to the three genotypes AA, AB, and BB of the tested SNP, $\nu$ was the effect of marker and treated as a fixed effect; Variable $X$ is an incidence matrix for non-genetic fixed effects, and $\beta$ is a non-genetic vector of fixed effects including month ages (time of birth to measurement), enter weight (weight of just entering the farm), fattening days (time of entering farm to measurement) and principal component effects (the top three eigenvectors of the $Q$ matrix). Variable $Z$ is an incidence matrix for a vector of polygenic effects, and parameter $u$ is a vector for residual polygenic effects with an assumed $N(0, K\sigma^2)$ distribution, where $\sigma^2$ is the polygenic variance and $K$ is a marker inferred kinship matrix. While $e$ is a vector for random residual errors with a putative $N(0, I\sigma_e^2)$ distribution, where $\sigma_e^2$ is the residual variance. The CMLM analysis was performed with GAPIT software package (http://www.maizegenetics.net/gapit). Quantile–quantile (Q–Q) plots were drawn to visualize the goodness of fitting for the GWAS model accounted by the population structure and familial relatedness. The negative logarithm of the $p$ value from the model was calculated against the expected value based on the null hypothesis. The threshold $p$ value after Bonferroni correction was $0.05/N = 7.45\times10^{-8}$, where $N$ is the number of SNPs. It’s too stringent. Such a Bonferroni correction results in a low statistical power [32]. Hence, we adopted the false discovery rate (FDR) to determine the threshold values for Single-GWAS, Multi-trait GWAS and LONG-GWAS. FDR was set as 0.01, and the threshold $p$ value was calculated as follows:

$$P = FDR \times n/m$$

where $n$ is the number of $P<0.01$ in the results, and $m$ is the total number of SNPs [33].

**Multi-trait GWAS**

The multi-trait GWAS were conducted to detect pleiotropic SNPs and the model was a Chi square statistic, which approximately follows a Chi square distribution with the number of traits tested as the number of degrees of freedom, was calculated for each SNP using the following formula [34]:

$$t_i = \frac{|\hat{v}_i|}{\sqrt{V(\hat{v}_i)}}$$
\[ \chi^2_{\text{multi-trait}} = t_i'V^{-1}t_i \]

Where \( \bar{\nu}_i \) is the estimate of \( v \) and the corresponding variance \( V(\bar{\nu}_i) \) can be obtained via compressed mixed linear model (CMLM). While \( t_i \) is a 6×1 vector of the signed t-values of the \( i \)th SNP from the above-mentioned single-trait GWAS for the six traits. Matrix \( t_i' \) is the transpose of the vector \( t_i \), and \( V^{-1} \) is the inverse of the 6×6 correlation matrix between traits, which was calculated by the estimated effects of the qualified SNPs (signed t values).

**LONG-GWAS**

As all experimental individuals were recorded for body size traits at three time stages, we conducted a longitudinal GWAS by LONG-GWAS [35]. This model was like CMLM except that the phenotypic variance was partitioned to SNPs, fixed factors (the above-mentioned \( \beta \) vector), polygenic effects, time stage effects and residual variance. Moreover, numerous studies have provided evidence that the longitudinal design could facilitate to identify time-dependent and consistent loci, which could increase the statistical power due to their effectiveness in incorporating the correlation structure of multiple measurements and alleviating the multiple testing burden [25, 26, 36]. While the code data implementing this method may be found at [http://genetics.cs.ucla.edu/longGWAS/](http://genetics.cs.ucla.edu/longGWAS/). This model is as follows:

\[ y^* = Wv + Zu + \gamma + e \]

In this formula, \( y^* \) is the adjusted phenotype. Variable \( W, Z, \) parameter \( v, u \) and \( e \) are consistent with CMLM mentioned above. Differently, the parameter \( \gamma \) is a vector for time stage effects with a putative \( N(0, \sigma^2_D) \) distribution, where \( D \) is a known block diagonal matrix representing the covariance between permanent environmental components. The \( D \) matrix was calculated by this formula: \( D = E \otimes I \), where \( E \) is a 3 × 3 matrix representing the covariance between the set of 3 time points for each individual.

**Abbreviations**

GWAS: Genome-wide association study; BH: body height; BL: body length; HH: hip height; HS: heart size; AS: abdominal size; CS: cannon bone size; SNP: single nucleotide polymorphism; QTL: quantitative trait loci; PCA: principal components analysis; CMLM: compressed mixed linear model;
Acknowledgements

Not applicable.

Authors’ contributions

BA wrote, and JL and HG revised the paper. LX1 and JX performed experiments. LZ and XG collected the GWAS data. TC, XW and LX2 interpreted the data. All authors reviewed and approved the final manuscript.

Funding

This work was supported by funds from the National Natural Science Foundations of China (NSFC, 31372294, 31472079 and 31702084) and Cattle Breeding Innovative Research Team (cxgc-ias-03). NSFC funded the collection, analysis, and interpretation of data. Cattle Breeding Innovative Research Team supported statistical analysis and writing of paper.

Availability of data and materials

We confirm that all raw data underlying our findings are publicly available without restriction. Data is available from the Dryad Digital Repository: doi:10.5061/dryad.4qc06.

Ethics approval and consent to participate

All animals used in the study were treated following the guidelines established by the Council of China Animal Welfare. Protocols of the experiments were approved by the Science Research Department of the Institute of Animal Sciences, Chinese Academy of Agricultural Sciences (CAAS) (Beijing, China). All individuals consented to the participation in this study. Written informed consent was obtained.

Consent for publication

Not applicable.
Competing interests

There are no competing interests.

Author details

1Institute of Animal Science, Chinese Academy of Agricultural Science, Beijing, 100193, China.

2Institute of Basic Medical Sciences, Westlake Institute for Advanced Study, Hangzhou, 310000, China.
References

1. Heinrichs AJ, Losinger WC. Growth of Holstein Dairy Heifers in the United States. J Anim Sci. 1998;76:1254–60.
2. Heinrichs AJ, Erb HN, Rogers GW, Cooper JB, Jones CM. Variability in Holstein heifer heart-girth measurements and comparison of prediction equations for live weight. Prev Vet Med. 2007;78:333–8.
3. Lund T, Miglior F, Dekkers JCM, Burnside EB. Genetic relationships between clinical mastitis, somatic cell count, and udder conformation in Danish Holsteins. Livest Prod Sci. 1994;39:243–51.
4. Vollema AR, Van Der Beek S, Harbers AGF, De Jong G. Genetic Evaluation for Longevity of Dutch Dairy Bulls. J Dairy Sci. 2000;83:2629–39.
5. 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;14.
6. Xia J, Qi X, Wu Y, Zhu B, Xu L, Zhang L, et al. Genome-wide association study identifies loci and candidate genes for meat quality traits in Simmental beef cattle. Mamm Genome. 2016;27:246–55.
7. Pryce JE, Hayes BJ, Bolormaa S, Goddard ME. Polymorphic regions affecting human height also control stature in cattle. Genetics. 2011;187:981–4.
8. Silventoinen K, Magnusson PKE, Tynelius P, Kaprio J, Rasmussen F. Heritability of body size and muscle strength in young adulthood: A study of one million Swedish men. Genet Epidemiol. 2008;32:341–9.
9. Zhang X, Chu Q, Guo G, Dong G, Li X, Zhang Q, et al. Genome-wide association studies identified multiple genetic loci for body size at four growth stages in Chinese Holstein cattle. PLoS One. 2017;12:e0175971.
10. Bouwman AC, Daetwyler HD, Chamberlain AJ, Ponce CH, Sargolzaei M, Schenkel FS, et al. Meta-analysis of genome-wide association studies for cattle stature identifies common genes that regulate body size in mammals. Nat Genet. 2018.
11. Lesosky M, Dumas S, Conradie I, Handel IG, Jennings A, Thumzi S, et al. A live weight-heart girth relationship for accurate dosing of east African shorthorn zebu cattle. Trop Anim Health Prod. 2012;45:311–6.
12. Dechow CD, Rogers GW, Klei L, Lawlor TJ. Heritabilities and Correlations Among Body
13. Dekkers JC. Commercial application of marker- and gene-assisted selection in livestock: strategies and lessons. Journal of animal science. 2004;82 E-Suppl.

14. Jiang L, Liu J, Sun D, Ma P, Ding X, Yu Y, et al. Genome wide association studies for milk production traits in Chinese Holstein population. PLoS One. 2010;5.

15. Meredith BK, Kearney FJ, Finlay EK, Bradley DG, Fahey AG, Berry DP, et al. Genome-wide associations for milk production and somatic cell score in Holstein-Friesian cattle in Ireland. BMC Genet. 2012;13.

16. 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;12.

17. Buzanskas ME, Grossi DA, Ventura R V., Schenkel FS, Sargolzaei M, Meirelles SLC, et al. Genome-wide association for growth traits in canchim beef cattle. PLoS One. 2014;9.

18. Jin B, Bao WJ, Wu ZQ, Xia XH. In situ monitoring of protein adsorption on a nanoparticulated gold film by attenuated total reflection surface-enhanced infrared absorption spectroscopy. Langmuir. 2012;28:9460–5.

19. Huang W, Kirkpatrick BW, Rosa GJM, Khatib H. A genome-wide association study using selective DNA pooling identifies candidate markers for fertility in Holstein cattle. Anim Genet. 2010;41:570–8.

20. Sahana G, Guldbrandtsen B, Bendixen C, Lund MS. Genome-wide association mapping for female fertility traits in Danish and Swedish Holstein cattle. Anim Genet. 2010;41:579–88.

21. Sorbolini S, Bongiorni S, Cellesi M, Gaspa G, Dimauro C, Valentini A, et al. Genome wide association study on beef production traits in Marchigiana cattle breed. J Anim Breed Genet. 2017;134:43–8.

22. Santana MHA, Utsunomiya YT, Neves HHR, Gomes RC, Garcia JF, Fukumasu H, et al. Genome-wide association analysis of feed intake and residual feed intake in Nellore cattle. BMC Genet. 2014;15.

23. Jahuey-Martínez FJ, Parra-Bracamonte GM, Sifuentes-Rincón AM, Martínez-González JC, Gondro C, García-Pérez CA, et al. Genomewide association analysis of growth traits in charolais beef cattle. J Anim Sci. 2016;94:4570–82.
24. Porter HF, O'Reilly PF. Multivariate simulation framework reveals performance of multi-trait GWAS methods. Sci Rep. 2017.

25. Das K, Li J, Wang Z, Tong C, Fu G, Li Y, et al. A dynamic model for genome-wide association studies. Hum Genet. 2011;129:629–39.

26. Kim S, Xing EP. Statistical estimation of correlated genome associations to a quantitative trait network. PLoS Genet. 2009;5.

27. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. Am J Hum Genet. 2007;81:559–75. doi:10.1086/519795.

28. Wang SB, Feng JY, Ren WL, Huang B, Zhou L, Wen YJ, et al. Improving power and accuracy of genome-wide association studies via a multi-locus mixed linear model methodology. Sci Rep. 2016;6.

29. Zhang Z, Ersoz E, Lai CQ, Todhunter RJ, Tiwari HK, Gore MA, et al. Mixed linear model approach adapted for genome-wide association studies. Nat Genet. 2010;42:355–60.

30. Lipka AE, Tian F, Wang Q, Peiffer J, Li M, Bradbury PJ, et al. GAPIT: Genome association and prediction integrated tool. Bioinformatics. 2012;28:2397–9.

31. VanRaden PM. Efficient Methods to Compute Genomic Predictions. J Dairy Sci. 2008;91:4414–23.

32. Nakagawa S. A farewell to Bonferroni: The problems of low statistical power and publication bias. Behavioral Ecology. 2004;15:1044–5.

33. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B. 1995;57:289–300.

34. Bolormaa S, Porto Neto LR, Zhang YD, Bunch RJ, Harrison BE, Goddard ME, et al. A genome-wide association study of meat and carcass traits in australian cattle. J Anim Sci. 2011.

35. Furlotte NA, Eskin E, Eyheramendy S. Genome-Wide Association Mapping With Longitudinal Data. Genet Epidemiol. 2012;36:463–71.

36. Zhou X, Stephens M. Efficient multivariate linear mixed model algorithms for genome-wide association studies. Nat Methods. 2014;11:407–9.

37. Arnold K, Sarkar A, Yram MA, Polo JM, Bronson R, Sengupta S, et al. Sox2 + adult stem and progenitor cells are important for tissue regeneration and survival of mice. Cell Stem Cell.
38. Yaman E, Gasper R, Koerner C, Wittinghofer A, Tazebay UH. RasGEF1A and RasGEF1B are guanine nucleotide exchange factors that discriminate between Rap GTP-binding proteins and mediate Rap2-specific nucleotide exchange. FEBS J. 2009;276:4607–16.

39. Chung J, Tsai S, James AH, Thames BH, Shytle S, Piedrahita JA. Lack of genomic imprinting of DNA primase, polypeptide 2 (PRIM2) in human term placenta and white blood cells. Epigenetics. 2012;7:429–31.

40. Haines BP, Rigby PWJ. Expression of the Lingo/LERN gene family during mouse embryogenesis. Gene Expr Patterns. 2008;8:79–86.

41. Guo Y, Huang Y, Hou L, Ma J, Chen C, Ai H, et al. Genome-wide detection of genetic markers associated with growth and fatness in four pig populations using four approaches. Genet Sel Evol. 2017;49.

42. Yi G, Shen M, Yuan J, Sun C, Duan Z, Qu L, et al. Genome-wide association study dissects genetic architecture underlying longitudinal egg weights in chickens. BMC Genomics. 2015;16.

43. Chen S, Li X, Lu D, Xu Y, Mou W, Wang L, et al. SOX2 regulates apoptosis through MAP4K4-Survivin signaling pathway in human lung cancer cells. Carcinogenesis. 2014;35:613–23.

44. Tani Y, Akiyama Y, Fukamachi H, Yanagihara K, Yuasa Y. Transcription factor SOX2 up-regulates stomach-specific pepsinogen A gene expression. J Cancer Res Clin Oncol. 2007;133:263–9.

45. Pausch H, Flisikowski K, Jung S, Emmerling R, Edel C, Götz KU, et al. Genome-wide association study identifies two major loci affecting calving ease and growth-related traits in cattle. Genetics. 2011;187:289–97.

46. Luechtefeld R. A general-purpose collaborative system: Theory and design of the dialogic web. 2017 IEEE Technol Eng Manag Soc Conf TEMSCON 2017. 2017;:455–9.

47. He M, Xu M, Zhang B, Liang J, Chen P, Lee JY, et al. Meta-analysis of genome-wide association studies of adult height in East Asians identifies 17 novel loci. Hum Mol Genet. 2015;24:1791–800.

48. Lundeen EA, Stein AD, Adair LS, Behrman JR, Bhargava SK, Dearden KA, et al. Height-for-age z scores increase despite increasing height deficits among children in 5 developing countries. Am J Clin Nutr. 2014;100:821–5.
49. Liu R, Sun Y, Zhao G, Wang H, Zheng M, Li P, et al. Identification of loci and genes for growth related traits from a genome-wide association study in a slow- × fast-growing broiler chicken cross. Genes and Genomics. 2015;37:829–36.

50. Hansen TH, Vestergaard H, Jørgensen T, Jørgensen ME ik., Lauritzen T, Brandslund I, et al. Impact of PTBP1 rs11085226 on glucose-stimulated insulin release in adult Danes. BMC Med Genet. 2015;16:17.

51. Meng Q, Wang K, Liu X, Zhou H, Xu L, Wang Z, et al. Identification of growth trait related genes in a Yorkshire purebred pig population by genome-wide association studies. Asian-Australasian J Anim Sci. 2017;30:462–9.

52. Dunn AR, Stout KA, Ozawa M, Lohr KM, Hoffman CA, Bernstein AI, et al. Synaptic vesicle glycoprotein 2C (SV2C) modulates dopamine release and is disrupted in Parkinson disease. Proc Natl Acad Sci. 2017;114:E2253–62.

53. Babette Z, Heidi K, Mingyao L, Sandra D, Celia K, Liming Q, et al. Role of bone-associated loci identified in GWAS meta-analyses in the context of longitudinal pediatric BMD in European Americans. Bone Abstr. 2013. doi:10.1530/boneabs.2.0c6.

54. Lee YS, Shin D, Song KD. Dominance effects of ion transport and ion transport regulator genes on the final weight and backfat thickness of Landrace pigs by dominance deviation analysis. Genes and Genomics. 2018;40:1331–8.

55. Du M, Auer PL, Jiao S, Haessler J, Altshuler D, Boerwinkle E, et al. Whole-exome imputation of sequence variants identified two novel alleles associated with adult body height in African Americans. Hum Mol Genet. 2014;23:6607–15.

56. Borowska A, Reyer H, Wimmers K, Varley PF, Szwaczkowski T. Detection of pig genome regions determining production traits using an information theory approach. Livest Sci. 2017;205:31–5.

57. Wang L, Zhang L, Yan H, Liu X, Li N, Liang J, et al. Genome-wide association studies identify the loci for 5 exterior traits in a large white × Minzhu pig population. PLoS One. 2014;9.

58. Klomp HM, Steyerberg EW, Ubbink DT. Letter: Systematic review and meta-analysis of controlled trials assessing spinal cord stimulation for inoperable critical leg ischaemia (Br J Surg 2004; 91: 948-955) [2] (multiple letters). Br J Surg. 2005;92:120–1.

59. Andersson L, Haley CS, Ellegren H, Knott SA, Johansson M, Andersson K, et al. Genetic
mapping of quantitative trait loci for growth and fatness in pigs. Science (80- ). 1994;263:1771–4.

60. Kemper KE, Reich CM, Bowman PJ, Vander Jagt CJ, Chamberlain AJ, Mason BA, et al. Improved precision of QTL mapping using a nonlinear Bayesian method in a multi-breed population leads to greater accuracy of across-breed genomic predictions. Genet Sel Evol. 2015;47.

61. Kemper KE, Littlejohn MD, Lopdell T, Hayes BJ, Bennett LE, Williams RP, et al. Leveraging genetically simple traits to identify small-effect variants for complex phenotypes. BMC Genomics. 2016;17.

62. Jung EJ, Park HB, Lee JB, Yoo CK, Kim BM, Kim HI, et al. Genome-wide association study identifies quantitative trait loci affecting hematological traits in an F2intercross between Landrace and Korean native pigs. Anim Genet. 2014;45:534–41.

63. Yariz KO, Walsh T, Uzak A, Spiliopoulos M, Duman D, Onalan G, et al. Inherited mutation of the luteinizing hormone/choriogonadotropin receptor (LHGR) in empty follicle syndrome. Fertil Steril. 2011;96:e125–30.

64. Zhang Z, Hong Y, Gao J, Xiao S, Ma J, Zhang W, et al. Genome-Wide Association Study Reveals Constant and Specific Loci for Hematological Traits at Three Time Stages in a White Duroc × Erhualian F 2 Resource Population. PLoS One. 2013;8.

65. Nishi H, Arai H, Momiyama T. NCI-H295R, a Human Adrenal Cortex-Derived Cell Line, Expresses Purinergic Receptors Linked to Ca2+-Mobilization/Influx and Cortisol Secretion. PLoS One. 2013;8.

66. Jia D, Jing Y, Zhang Z, Liu L, Ding J, Zhao F, et al. Amplification of MPZL1/PZR promotes tumor cell migration through Src-mediated phosphorylation of cortactin in hepatocellular carcinoma. Cell Res. 2014;24:204–17.

67. Rask-Andersen M, Almén MS ällma., Lind L, Schiöth HB. Association of the LINGO2-related SNP rs10968576 with body mass in a cohort of elderly Swedes. Mol Genet Genomics. 2015;290:1485–91.

68. Bryant EK, Dressen AS, Bunker CH, Hokanson JE, Hamman RF, Kamboh MI, et al. A Multiethnic Replication Study of Plasma Lipoprotein Levels-Associated SNPs Identified in Recent GWAS. PLoS One. 2013;8.

69. Guo YM, Zhang ZY, Ma JW, Ai HS, Ren J, Huang LS. A genomewide association study of feed efficiency and feeding behaviors at two fattening stages in a white duroc × erhualian F2 population. J
Anim Sci. 2015;93:1481–9.
Figure 1. Principal components (PC) plot drawn from the second principal component (PC2) against the first principal component (PC1).

Figure 2, Figure 3. The strengths of genome-wide association studies (GWAS) are illustrated by the Manhattan plots on the left panel. The deviations of the signals from null hypothesis are illustrated as the Quantile-Quantile (QQ) plots on the right panel. The negative logarithms of the observed (y axis) and the expected (x axis) P values are plotted for each SNP (dot). GWAS were performed six body size traits months 6, 12 and 18 after birth separately. Each analysis is labeled as trait (BH or HH) and month on the far right. The number neighboring each trait indicates the age of measurement (e.g., BH6 = Body Height at 6 months). The 29 chromosomes are color coded.