

Insight Into the Causative Genes and Pathways Associated With Height, Length, Length to Height Ratio and Body Weight Traits of Korean Indigenous Breed, Jindo Dog Using Gene Set Enrichment and Pathway-based GWAS Analysis

Sunirmal Sheet

National Institute of Animal Science

Jong Seok Kim

Korean Jindo and Domestic Animals Center

Min Jeong Ko

National Institute of Animal Science

Na Yeon Kim

National Institute of Animal Science

Young Jo Lim

National Institute of Animal Science

Mi Rim Park

National Institute of Animal Science

Seung Jin Lee

Korean Jindo and Domestic Animals Center

Jeong Min Kim

Korean Jindo and Domestic Animals Center

Seok Il Oh

Korean Jindo and Domestic Animals Center

Bong-Hwan Choi (✉ bhchoi@korea.kr)

National Institute of Animal Science

Research Article

Keywords: Indigenous breed, Jindo, quantitative trait, GWAS, gene-set enrichment, heritability

Posted Date: July 26th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-738677/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

As a companion and hunting dog, height, length, length to height ratio (LHR) and body-weight are the vital economic traits for Jindo dog. Artificial breeding has produced an extraordinary diversity in these traits. Therefore, the identification of causative markers, genes and pathways that led us to understand the genetic basis of this variability is essential for their selection purposes. Here, we performed a genome-wide association study (GWAS) combined with pathway-based analysis on 757 dogs using 118,879 SNPs. A higher heritability (h^2) was detected for height (0.33) and weight (0.28) trait in Jindo. At a threshold of p -value $< 5E-05$, 10, 6, 13, and 11 SNPs on different chromosomes were significantly associated with height, length, LHR and body-weight traits, respectively. Based on our result, *HHIP*, *LCORL*, and *NCAPG* for height, *IGFI* and *FGFR3* for length, *DLK1* and *EFEMP1* for LHR, and *PTPN2*, *IGFI*, and *RASAL2* for weight can be the potential candidate genes because the significant SNPs located in their intronic or upstream regions. An additive and dominant mode of inheritance was noticed from the phenotype-genotype correlation plot for top variants. The gene-set enrichment analysis highlighted here 9 and 7 overlapping significant ($p < 0.05$) GO terms and pathways among traits. Interestingly, the highlighted pathways were related to hormone synthesis, secretion and signaling were generally involved in the metabolism, growth and development process. Our data provide an insight into the significant genes and pathways if verified further, which will have a significant effect on the breeding /management of the Jindo dog's population.

Introduction

The Jindo dog is a magnificent Korean indigenous breed that originated from the Jindo county of South Korea. It has been highly regarded for its hostile disposition and aggressive characteristics. It is well known as a loyal companion, independent hunters and discerning guardians. In 1998, the Jindo dog breed got recognition from United Kennel Club and Federation (Kim et al. 2012). This breed has dark brown bright coloured eyes, small erected triangular ears, and rolled up or sickle-shaped tails with acute hearing and sense of smell (Kim et al. 2010b). Jindo dogs are a double-coated spitz-type dog with ranges from 48 to 53 cm in height, and the body-weight range is 18 to 27 kg. The majority of Jindo owners keep them for farmhouse income, watchdog and hunting. These days, the dogs are intensively farmed and kept as affectionate companion dog. The Jindo dog ranked 8th in the list of canine breeds kept as pets, falling behind other foreign breeds, and was found in only 4% of households that had a domestic dog (Kim et al. 2009).

Throughout the journey to domestication, the dog has become a part of the different functional and emotional needs of humans (Stone et al. 2016). Domestic dogs nowadays are descended from random breeding or purpose-bred populations in which the selection of their parents is under humans control (Bannasch et al. 2020; Pilot et al. 2015). The intense artificial selection has resulted in a diverse range of dogs that serve as herders, hunters, protectors, trackers, support dogs, athletes, and, most popularly, as companions (Stone et al. 2016). Specifically, the purpose-bred dogs were produced for performing different highly specialized tasks (Albarella 2003; Bannasch et al. 2020; Lazarowski et al. 2018). For example, certain dogs were used to carry or pull bulky loads and consequently were selected for larger overall size and bones (Bannasch et al. 2020; Club 1998). Some dogs were designated for speed to hunt or herd game and thus have longer legs and leaner builds (Bannasch et al. 2020; Club 1998; Spady and Ostrander 2008; Young and Bannasch 2006). Besides these functional purposes, various breeds were produced exclusively as companion animals (Bannasch et al. 2020; Club 1998). In such circumstances, morphology was subjected to fewer restrictions, allowing for the selection of more extremes as required. Body sizes in domestic dogs can vary up to 50-times, from toy breeds to giant breeds.

One of the most visual examples of artificial selection for domestic dogs is body size variations including changes in height, length, and body weight across (Kim et al. 2010b; Lu et al. 2020; Plassais et al. 2019; Plassais et al. 2017). The height, weight and other characteristics of the body's size are the classic polygenic and highly heritable morphologic traits, influenced by several variants across different genetic loci (Allen et al. 2010; Chan et al. 2015; Plassais et al. 2019; Rimbault et al. 2013). Therefore, studies on a better understanding of the genes and variants that affect morphological variation in domestic dogs can be found from the late 19th century by Charles Stockard in (Stockard 1941) to today by modern canine geneticists.

A genome-wide association study (GWAS) is a powerful approach used in genetic research for identifying causal gene and loci associated with complex diseases and quantitative traits (Srikanth et al. 2020). In recent years, the decreasing cost of high-throughput genotyping technology integrated with the ongoing gene set enrichment analysis (GSEA) has led to a boom in the number of genome-wide association studies. GWAS in companion animal, including the dog, have provided significant insights into susceptible genes and genomic loci associated with numerous canine disease traits and many Mendelian disorder-related traits (Drögemüller et al. 2014; Momozawa et al. 2020; Sheet et al. 2020; Signer-Hasler et al. 2012). In addition to complex disease traits, GWAS studies have also been used to discover genes affecting different polygenic morphological traits in dogs as well as many other mammals and, of course,

humans (Allen et al. 2010; Boyko et al. 2010; Dadousis et al. 2017; Hayward et al. 2016; Signer-Hasler et al. 2012; Takasuga 2016). For example, a total of 13 loci have been discovered in dogs that have a significant impact on their weight and/or height (Boyko et al. 2010; Chase et al. 2002; Eigenmann et al. 1984; Hoopes et al. 2012; Jones et al. 2008; Parker et al. 2009; Plassais et al. 2019; Quignon et al. 2009; Rimbault et al. 2013; Sutter et al. 2007; Vaysse et al. 2011). Among the 13 loci, six loci have been reported for accounting for more than 80% of the diversity in body size trait in purebred dogs (Boyko et al. 2010; Hayward et al. 2016). On the other hand, GWAS in humans based on tens of thousands to hundreds of thousands of samples so far revealed ~ 700 loci for human height explaining only 20 % of the variation (Allen et al. 2010; Signer-Hasler et al. 2012).

In this study, we performed a GWAS analysis to investigate the candidate genes and biological pathways related to quantitative traits (height, length, length to height ratio (LHR), and body-weight) of 757 Korean Jindo dogs using SNP genotyping data from Illumina CanineHD BeadChip array and supplemented it with gene-set and pathway-based functional analysis. The significantly associated SNPs, candidate genes, gene ontology (GO) and pathways related to height, length, LHR, and body weight were detected and thus, giving references for use during the marker-assisted selection of these traits in Jindo dog.

Materials And Methods

Animals and Phenotype assignment

Institutional Animal Care and Use Committee (IACUC) approval at the National Institute of Animal Science (NIAS, RDA), was obtained for this study. The protocol consent was obtained for the 'Development of early diagnosis technology for degenerative muscular skeleton system in special-purpose dog' project. A total of 757 Jindo dogs were selected from the farm of Jindo county of South Korea in the present study. The dogs were selected randomly for this study, without pedigree information. All dogs were selected without obesity-related diseases which could affect body weight and kept under fasting for 12 h from their last meal to get the dog's true weight in the morning. The blood sample was collected at the same time. A dog's height measurement was taken from the withers' highest point down to the paws. The length measurement was taken from starting at the centre of the chest along the side of the body to the tail. The height and length were measured in centimetre (cm) and weight was measured in kilogram (kg). The following formula was used to get the LHR -

$$\text{LHR} = \left(\frac{\text{Length}}{\text{Height}} \right) \times 100$$

1

The other relevant information (e.g., identification number, intake date, age, sex, and obesity-related) was retrieved from shelter records.

Genomic DNA extraction, SNP Genotyping and quality control

We extracted genomic DNA from collected blood samples of 757 dogs, using the DNeasy Blood and Tissue Kit (Qiagen, Valencia, CA, USA). The samples were genotyped on a 170K Illumina CanineHD BeadChip (Illumina, San Diego, CA, USA) array, which contains 173,662 SNPs. The filtering criteria were set as follows: minor allele frequency < 5%, low genotyping call rate < 90 %, missing genotype calls > 10 %, Hardy–Weinberg equilibrium at $p < 1E-06$ for quality control (QC) with PLINK v.1.9 software [Srikanth et al., 2020] (Srikanth et al. 2020). The final genotyping call rate was 98.75 %. After QC filtering, 118,879 SNPs and 757 animals remained for further association analysis.

Genome-wide association analysis

The association between traits and SNPs were tested using a mixed linear model approach implemented in GCTA v.191.4 beta3. Significant factors such as age (1-12), sex (492 females and 265 males), and birth year (2006-2019) were fitted in the GWAS statistical model as a fixed effect for all the traits. In GWAS, we generated a total of 20 principal components (PCs); the eigenvalues of all the PCs were fit as co-variance to account for population stratification. The GWAS statistical model used was as follows:

$$y = a + bx + g + e \quad (2)$$

where y is a phenotypic trait, a is mean, b is an additive effect (fixed effect), u is additive genetic effect for each marker, x is the matrix of SNPs coded as 0, 1, or 2, g is the accumulated effect of all the SNPs captured by the GRM (genetic relationship matrix, calculated using all the SNPs), and e is a vector of residual (Srikanth et al. 2020). The significance threshold for GWAS was distinct using the

Bonferroni correction method. A Bonferroni corrected threshold was used to correct for multiple testing, and the genome-wide significance threshold at 5% was $p < 4E-07$ ($\sim 0.05/118,879$). However, the Bonferroni-corrected threshold was too stringent in this study and therefore, it might yield many false-negative results. Hence, the variants with p nominal value $< 5E-05$ were considered genome-wide significant to reduce several false-negative results caused by the Bonferroni-correction threshold (Do et al. 2017). Manhattan and quantile-quantile (Q-Q) plots were generated using the CMplot package in R (Srikanth et al. 2020). The estimate of variance ($V(G)$ and V_p) was performed using restricted maximum likelihood analysis (REML) implemented in GCTA v.191.4. beta3, while heritability (h^2) was calculated using $h^2 = V(G)/V_p$ (Schmidt et al. 2019).

Gene mapping, GO and pathway analysis

We performed a gene-set enrichment and pathway analysis for each trait following the methods described by Dadousis et al (Dadousis et al. 2017). We used a nominal p -value of < 0.01 to filter for SNPs from the GWAS for gene-set and pathway analysis. If the annotation of genes using only significant SNPs carried out, it may not encounter the SNPs with less significant level. As a result, it will miss key putative genes and allied pathways (Do et al. 2017). Moreover, it has been shown that merging less significant but associated SNPs can provide information on how these markers might be collectively related to our phenotypes of interest (Sheet et al. 2020). The SNPs were annotated to genes within 5-kb flanking region using SnpEff version 4.3 software (Srikanth et al. 2020) and the genes were used in the enrichment analysis. Genes name assigned to SNPs was filtered using SNP ID's from variant call format (VCF) file that was used for mapping. The associated genes were then assigned to Database for Annotation, Visualization, and integrated discovery (DAVID, <http://david.abcc.ncifcrf.gov/>) for functional annotation and enrichment analyses (Srikanth et al. 2020). The Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database and GO database under biological process and molecular function categories were used (Lee et al. 2019b). Over-representation of genes in every GO terms and pathways were tested using Fisher's exact test, which was conducted using the DAVID database (Lee et al. 2019a; Lee et al. 2019b). False discovery rate (FDR) correction (≤ 0.05) was applied to account for multiple testing (Lee et al. 2019a). Significant enrichment of pathways and GO terms were indicated by a p -value of ≤ 0.05 .

Result

Phenotypic Data Analysis

We performed a GWAS with different quantitative traits such as height, length, LHR, and body-weight in Jindo dog. We investigated 757 Jindo dogs with an average raw height of 48.51 cm, ranging from 33.5 to 59 cm and an average length of 51.42 cm ranging from 37 cm to 62 cm. The average LHR is 105.82, ranging from 86.4 to 130.7. The average raw weight in this individual dataset is 16.67 kg, ranging from 6.6 kg to 33.3 kg.

The size variance was noticed high because of different ages of dogs were included in our current study for example a medium dog (younger than 16 months), adult dog (up to 8 years), and senior dogs (8 years or older). The standard error of the mean value was 0.13, 0.13, 0.18, and 0.16 for height, length, LHR, and body weight, respectively. The distribution of dogs was done according to bodyweight quartiles such as Q1 (the lowest 25 % of registered weight) – 8.4, Q2 and Q3–25 % below (12.4) and above the median (20.8), Q4 – dogs with the highest 25% of registered weight values (24.6). A total of 9 (Q1), 46 (Q2), 603 (Q3), and 99 (Q4) dogs were in each body-weight quartile. The ratios of genotypic and phenotypic variance explained by all the SNPs were found to be 0.33, 0.08, 0.08, and 0.28 for height, length, LHR, and body weight, respectively. Figure 1 shows an example of phenotype measurement of male and female Jindo dogs. A descriptive statistical summary of the phenotypes included the minimum, maximum, and mean, as well as the standard deviations and standard error are given in Table 1.

Table 1
Descriptive summary of phenotype data.

Traits	Phenotypic Data						Variance		
	492 Female, 265 Male								
	Mean	Median	Max	Min	SD	SE	h ²	Vp	V(G)
Height (cm)	48.61	48.5	59	33.5	3.69	0.13	0.33	13.26	4.31
Length (cm)	51.42	51.3	62	37	3.62	0.13	0.08	106.68	8.81
LHR	105.82	105.6	130.7	86.4	5.23	0.18	0.08	361.40	32.14
Weight (Kg)	16.67	16.4	33.3	6.6	3.41	0.16	0.28	11.34	3.19

SD – standard deviation, SE – standard error, Vp – phenotypic variance, V(G) – genotypic variance, h² – heritability, LHR – Length to height ratio.

Genome-wide Association Study

Before carrying out the GWAS analysis, we analyzed the patterns of population structure through principal component analysis (PCA) and the population found to form a good clustering pattern (Fig. 2). We performed GWAS with a total of 757 dogs and 118,879 SNPs that passed after QC to discover significant loci associated with traits. The total genotyping rate was 0.988285. The Manhattan plot represents our mixed linear model-based GWAS result (Figs. 3, 4, 5 and 6). The association result revealed that a total of 40 SNPs in the genome were associated with all the four traits of height, length, LHR and body-weight at a genome-wide significant level ($p < 5E-05$). Among these SNPs, 10, 6, 13, and 11 SNPs were significant at the genome-wide level ($p < 5E-05$) for the trait of height, length, LHR and body-weight, respectively. A strong signal was identified only for height and body-weight trait, where one SNP (TIGRP2P201033_rs9187576) on Chr15 and 4 SNPs on Chr7 and Chr15 passed the Bonferroni corrected significance threshold ($p < 4E-07$). Most of these SNPs were located in the intronic regions and intergenic regions. No SNPs crossed Bonferroni corrected significance threshold ($p < 4E-07$) for the trait of length and LHR. The topmost significant ($p < 5E-05$) five associated SNPs and their annotated genes are presented in Table 2. For the height trait, top five SNPs (TIGRP2P201033_rs9187576, BICF2G630358507, BICF2G630358640, BICF2S23724992, and BICF2S2377250) were annotated to four genes (*HHIP*, *LCORL*, *NCAPG*, and *P4HA1*), respectively. For the length trait, the SNPs (BICF2P772349, BICF2P580549, BICF2S2366810, BICF2P517149, and BICF2G630778876) were annotated to four different genes such as *FGFR3*, *PLCH1*, *IGF1*, and *SCAND3*, respectively. For LHR, the SNPs (TIGRP2P119932_rs8658799, BICF2G630268912, BICF2G630490347, BICF2P491431, and BICF2G630507820) were successfully annotated to seven genes (*BEGAIN-DLK1*, *TDRD1*, *EFEMP1*, *ZFH3-PSMD7*, and *SLC23A2*) genes, respectively. For body-weight, the SNPs (BICF2P279062, BICF2P979272, BICF2S23655947, BICF2S2336786, and BICF2P110929) were successfully annotated to seven genes (*PTPN2*, *RASAL2*, *PARPBP-IGF1*, *IGF1*, *ENSCAFG00000001744-SNORD19*) genes, respectively. The clear deviation between observed and expected p -values on the Q-Q plot for all the traits signifies good correlation results in the present study (Figs. 3, 4, 5 and 6).

Table 2
Top 5 SNPs associated with height, length, LHR, and body weight in Jindo dog.

Trait	SNP ID	Chr	Position	p-value	Freq	Gene	Type
Height	TIGRP2P201033_rs9187576	15	43493060	1.49E-07	0.11	<i>HHIP</i>	intron_variant
	BICF2G630358507	3	91180492	6.91E-07	0.47	<i>LCORL</i>	intron_variant
	BICF2G630358640	3	91314456	2.1E-06	0.33	<i>NCAPG</i>	intron_variant
	BICF2S23724992	4	23625990	4E-06	0.10	<i>P4HA1</i>	intron_variant
	BICF2S2377250	3	93890772	6.65E-06	0.11	-	-
Length	BICF2P772349	3	62323843	4.42E-06	0.29	<i>FGFR3</i>	Upstream_gene_variant
	BICF2P580549	23	49471629	5.42E-06	0.27	<i>PLCH1</i>	Intron_variant
	BICF2S2366810	31	40675954	3.71E-05	0.25	-	-
	BICF2P517149	15	41228320	3.98E-05	0.24	<i>IGF1</i>	Intron_variant
	BICF2G630778876	35	25711203	4.18E-05	0.17	<i>SCAND3</i>	Intron_variant
LHR	TIGRP2P119932_rs8658799	8	68874072	7.88E-07	0.36	<i>BEGAIN-DLK1</i>	Intron_variant
	BICF2G630268912	28	25034623	8.78E-07	0.28	<i>TDRD1</i>	Intron_variant
	BICF2G630490347	10	56634077	8.98E-07	0.39	<i>EFEMP1</i>	Intron_variant
	BICF2P491431	5	79708771	1.20E-06	0.11	<i>ZFH3-PSMD7</i>	Intergenic_region
	BICF2G630507820	24	16702908	4.09E-06	0.07	<i>SLC23A2</i>	Intron_variant
Weight	BICF2P279062	7	78202038	8.91E-08	0.03	<i>PTPN2</i>	intron_variant
	BICF2P979272	7	21586876	1.54E-07	0.37	<i>RASAL2</i>	intron_variant
	BICF2S23655947	15	41134849	2.12E-07	0.05	<i>PARPBP-IGF1</i>	intergenic_region
	BICF2S2336786	15	41239301	2.86E-07	0.07	<i>IGF1</i>	intron_variant
	BICF2P110929	14	10389714	1.06E-06	0.03	<i>ENSCAFG00000001744-SNORD19</i>	intergenic_region
LHR - Length to height ration							

Genotype-phenotype Correlation

We acquired height, length, LHR, and body-weight measurements on genotyped Jindo dogs.

We then compared the traits (height, length, LHR, and body-weight) distributions between the different genotype classes at the topmost variants, TIGRP2P201033_rs9187576, BICF2P772349, TIGRP2P119932_rs8658799, and BICF2P279062, respectively. The box plots

show the distribution of phenotypes among the different genotypes (Fig. 7). The solid lines in the box plot are the medians of phenotype's per-genotype group (AA, AG, GG, CC, CG, and CG) of all four variants were described in Table 3. In addition, we calculated the mean height, length, LHR, and body weight per genotype (Table 3). Although the differences between all three genotype classes were not significant, it showed that the A-allele for SNP TIGRP2P201033_rs9187576 was correlated with a mean reduction of the height trait in Jindo. For SNP BICF2P772349 and TIGRP2P119932_rs8658799, the A-allele showed to be correlated with increased length and LHR in Jindo dog, respectively. The trait-increasing allele A for SNP TIGRP2P119932_rs8658799 and G for SNP TIGRP2P201033_rs9187576 was found to be a major allele for both the SNPs. This A allele was detected as a minor allele in the case of BICF2P772349 SNP of length trait. However, the AA genotype for SNP BICF2P772349 was found to be present only in a small number of total animals.

Table 3
Effect of top SNP genotypes on the height, length, LHR, and weight trait in Jindo dog.

Traits	SNP name	Genotype	Counts	Mean	Median	SD
Height (cm)	TIGRP2P201033_rs9187576	AA	16	48.1	48.1	3.96
		AG	144	48.37	48.50	3.82
		GG	597	48.70	49	3.60
Length (cm)	BICF2P772349	AA	2	53.26	53.00	-
		AG	126	51.08	52.00	3.49
		GG	629	51.44	52.00	3.64
LHR	TIGRP2P119932_rs8658799	AA	306	105.93	104.90	4.98
		AG	349	105.69	104.86	4.32
		GG	102	105.67	104.85	5.86
Weight (kg)	BICF2P279062	CC	95	16.14	16.05	3.36
		CG	335	16.96	17	3.72
		GG	327	16.53	16	3.08
LHR – Length to height ration						

Gene-set Enrichment And Pathway Analysis

Functional enrichment analysis was carried out to identify classes of genes that are over-represented in a large group of genes and may have a connection with the studied phenotypes. Many genes are in GO term and KEGG pathway categories. From the GWAS result, 1222, 1132, 1102 and 1039 SNPs (nominal $p < 0.01$) were used to annotate 842, 817, 752, and 718 genes correlated with height, length, LHR and weight trait (Supplementary Table S1). Subsequently, the genes were uploaded to the bioinformatics tool, DAVID for finding the enriched GO terms and KEGG pathways. The analysis revealed that a total of 72 GO terms and 37 KEGG pathways for height, 69 GO terms and 19 KEGG pathways for length, 53 GO terms and 18 KEGG pathways for LHR, 50 GO terms and 25 KEGG pathways for weight traits were significantly ($p < 0.05$) enriched. Out of the total enriched GO terms and KEGG pathways, the top 5 significantly enriched GO terms and pathways are presented in Table 4. For height, the top 5 GO terms were GO:1902711 – GABA-A receptor complex, GO:0004890 – GABA-A receptor activity, GO:0006816 – Calcium ion transport, GO:0030336 – Negative regulation of cell migration, and GO:0045665 – Negative regulation of neuron differentiation, and KEGG pathways were cfa04024 – cAMP signaling pathway, cfa04723 – Retrograde endocannabinoid signaling, cfa04261 – Adrenergic signaling in cardiomyocytes, cfa05033 – Nicotine addiction, and cfa04540 – Gap junction. For length, the top 5 GO terms were GO:0030425 – Dendrite, GO:0005096 – Gtpase activator activity, GO:0043025 – Neuronal cell body, GO:0032332 – Positive regulation of chondrocyte differentiation, and GO:0042593 – Glucose homeostasis, and KEGG pathways were cfa04922 – Glucagon signaling pathway, cfa04915 – Estrogen signaling pathway, cfa04024 – cAMP signaling pathway, cfa04918 – Thyroid hormone synthesis, and cfa04080 – Neuroactive ligand-receptor interaction. The top 5 GO terms (GO:0048706 – Embryonic skeletal system development, GO:0010595 – Positive regulation of endothelial cell migration, GO:0050840 – Extracellular matrix binding, GO:0048484 – Enteric nervous system development, and GO:1901166 – Neural crest cell migration involved in autonomic nervous system development) and KEGG pathways (cfa04911 – Insulin secretion, cfa04080 –

Neuroactive ligand-receptor interaction, cfa04931 – Insulin resistance, cfa04915 – Estrogen signaling pathway, and cfa05202 – Transcriptional misregulation in cancer) were significantly enriched for LHR. For weight, the top 5 GO terms were GO:0030425 – Dendrite, GO:0005096 – Gtpase activator activity, GO:0017137 – Rab gtpase binding, GO:0045666 – Positive regulation of neuron differentiation, and GO:0050852 – T cell receptor signaling pathway, and KEGG pathways were cfa04921 – Oxytocin signaling pathway, cfa05100 – Bacterial invasion of epithelial cells, cfa04713 – Circadian entrainment, cfa04918 – Thyroid hormone synthesis, and cfa04520 – Adherens junction.

Table 4

Top 5 Gene Ontology and KEGG pathways significantly enriched using genes associated with height, length, LHR and body weight.

Trait	Category	Term_ID	Term	Count	%	P-value	Genes
Height	GOTERM_CC_DIRECT	GO:1902711	GABA-A receptor complex	6	0.004585	6.E-04	<i>GABRA2, GABRB2, GABRA6, GABRA5, GABRA4, GABRG1</i>
	GOTERM_MF_DIRECT	GO:0004890	GABA-A receptor activity	6	0.004585	7.E-04	<i>GABRA2, GABRB2, GABRA6, GABRA5, GABRA4, GABRG1</i>
	GOTERM_BP_DIRECT	GO:0006816	Calcium ion transport	8	0.006114	8.E-04	<i>PPP3CA, RYR2, CAMK2D, PLN, PRKCB, CCL5, CAV1, NMUR2</i>
	GOTERM_BP_DIRECT	GO:0030336	Negative regulation of cell migration	11	0.008407	9.E-04	<i>RNF20, KANK1, DACH1, STK24, DAG1, ADARB1, LDLRAD4, TBX5, SULF1, RECK, ROBO1</i>
	GOTERM_BP_DIRECT	GO:0045665	Negative regulation of neuron differentiation	9	0.006878	9.E-04	<i>DDX6, ZHX2, MEIS1, ID4, OLIG2, CNTN4, PBX1, GLI3, CDK5RAP2</i>
	KEGG_PATHWAY	cfa04024	Camp signaling pathway	23	0.017577	6.E-06	<i>GRIA1, RYR2, GABBR2, CAMK2D, CREBBP, SUCNR1, PPP1R12A, BDNF, HHIP, ADCY2, ADRB1, ATP2B1, GLI3, GNAI1, MAPK10, TIAM1, GRIN3A, PLN, FXYD2, CREB3L2, DRD1, CREB5, GRIA4</i>
	KEGG_PATHWAY	cfa04723	Retrograde endocannabinoid signaling	16	0.012228	1.E-05	<i>GABRA2, GRIA1, GABRB2, GABRA6, GABRA5, PRKCB, GABRA4, ADCY2, GNAI1, GRM1, GABRG1, MAPK10, RIMS1, GRM5, PLCB1, GRIA4</i>
	KEGG_PATHWAY	cfa04261	Adrenergic signaling in cardiomyocytes	17	0.012992	8.E-05	<i>RYR2, CAMK2D, PPP2R5A, ADCY2, ADRB1, ATP2B1, ADRA1B, GNAI1, PLN, PPP2R2B, PPP2R5E, KCNQ1, FXYD2, CREB3L2, PLCB1, CACNG3, CREB5</i>
	KEGG_PATHWAY	cfa05033	Nicotine addiction	9	0.006878	1.E-04	<i>GABRA2, GRIA1, GABRB2, GRIN3A, GABRA6, GABRA5, GABRA4, GRIA4, GABRG1</i>
	KEGG_PATHWAY	cfa04540	Gap junction	13	0.009935	2.E-04	<i>GUCY1A3, GUCY1B3, GUCY1A2, PRKCB, LPAR1, ADCY2, ADRB1, GNAI1, GRM1, TUBB6, GRM5, DRD1, PLCB1</i>

Table 4
Cont.

Trait	Category	Term_ID	Term	Count	%	P value	Genes
Length	GOTERM_CC_DIRECT	GO:0030425	Dendrite	21	2.781457	8.E-07	<i>GABRA2, EPHA4, SYT4, GABRA5, KCNB1, DENND1A, MAGI2, KLHL1, PDYN, GRM1, BRINP2, HTR7, RELN, PLCB4, BRINP1, GRM7, DNER, NCS1, MAPT, BRINP3, TP63</i>
	GOTERM_MF_DIRECT	GO:0005096	Gtpase activator activity	21	2.781457	8.E-05	<i>RGS18, RABGAP1, DAB2IP, ARHGAP18, RASAL2, ARFGAP3, ARHGAP15, ARHGAP24, TIAM2, TBC1D1, TBCK, TBC1D2, SIPA1L1, TBC1D21, DLC1, TBC1D13, RAPGEF2, PLCB1, DOCK1, GARNL3, RGS7</i>
	GOTERM_CC_DIRECT	GO:0043025	Neuronal cell body	15	1.986755	7.E-04	<i>GABRA2, DENND1A, SEZ6, KLHL1, GRIK2, PDYN, BRINP2, UCHL1, GRIN3A, BRINP1, DNER, RAPGEF2, APOB, ASIC2, BRINP3</i>
	GOTERM_BP_DIRECT	GO:0032332	Positive regulation of chondrocyte differentiation	5	0.662252	2.E-03	<i>PKDCC, SOX9, SOX6, GLI3, SOX5</i>
	GOTERM_BP_DIRECT	GO:0042593	Glucose homeostasis	3	1.675676	2.E-03	<i>TFAP2B, RFX6, MCU</i>
	KEGG_PATHWAY	cfa04922	Glucagon signaling pathway	6	1.6275	3.E-02	<i>PLCH1, ATP6V1G3, ACACA, CREB3L2, SLC2A2, CREB5</i>
	KEGG_PATHWAY	cfa04915	Estrogen signaling pathway	4	2.385	2.E-03	<i>CREB3L2, ADCY8, GRM1, CREB5</i>
	KEGG_PATHWAY	cfa04024	cAMP signaling pathway	20	2.649007	1.E-04	<i>ABCC4, BDNF, CACNA1D, ADRB1, ADCY8, GRIN2B, GLI3, GNAI1, PPP1CC, TIAM1, GRIN3A, PLN, FXD2, ORAI1, PLCE1, FFAR2, SOX9, CNGA4, GRIA4, RAPGEF4</i>
	KEGG_PATHWAY	cfa04918	Thyroid hormone synthesis	3	2.6375	3.E-02	<i>CREB5, PLCH1, SCAND3</i>
	KEGG_PATHWAY	cfa04080	Neuroactive ligand-receptor interaction	17	2.251656	4.E-02	<i>GABRA2, GABRB2, GRID1, GABRA5, ADRB1, TACR1, NMBR, GRIK2, PRL, GRIN2B, GRM1, GABRG1, GRIN3A, HTR7, GRM4, GRM7, GRIA4</i>

Moreover, we have highlighted here the significant enrichment pathways and GO terms that were overlapping among the traits. Among the total enriched GO terms and KEGG pathways, significantly ($p < 0.05$) enriched 9 GO terms and 7 pathways were found to be shared among traits (Fig. 8). The enriched GO terms are GO:0005634 - nucleus, GO:0047497 - mitochondrion transport along the microtubule, GO:0051569 - Thyroid hormone synthesis GO:0045595 - regulation of cell differentiation, GO:0051561 - positive regulation of mitochondrial calcium ion concentration, GO:0042593 - glucose homeostasis, GO:0050840 - extracellular matrix binding, GO:0015629 - actin cytoskeleton, GO:0005829 - cytosol. The KEGG pathways are cfa04911 - Insulin secretion, cfa04915 - Estrogen signaling pathway, cfa04925 - Aldosterone synthesis and secretion, cfa04922 - Glucagon signaling pathway, cfa04931 - Insulin resistance, cfa04918 - Thyroid hormone synthesis, cfa04022 - cGMP-PKG signaling pathway.

Discussion

The present study reveals previously unreported information of genetic contribution to the important morphological phenotypes such as height, length, LHR and body weight of Korean native breed, Jindo dog. Growth-related trait such as height, length, and body weight are reported to have a higher heritability percentage (Helmink et al. 2001). Therefore, a higher heritable trait always needs to be precisely selected. Moreover, it has been shown in a survey study on the breeds' height, bodyweight and related to data on 36 behavioural traits of companion dogs (n = 8,301) of various common breeds (n = 49) collected internationally using the Canine Behavioral Assessment and Research Questionnaire (C-BARQ) that particular canine morphotypes tend to be reliably allied with particular behavioural profiles (McGreevy et al. 2013). Here, marker-assisted selection can also be used to predict an animal's phenotypic potential using molecular genetics (Bray et al. 2021). Therefore in this study, we carried out GWAS supplemented with gene-set enrichment analysis to identify important functional candidate genes and pathways that influence the morphological phenotypes such as height, length, LHR and body-weight traits in Jindo dog and understand their regulatory role.

We undertook an association study of 173,662 variants in 757 individual dogs and identified a total of 40 SNPs associated with height, length, LHR, and body-weight traits at the genome-wide level ($p < 5E-05$). A mixed-linear model has been used in this study to control the population stratification effect as it is the most effective method which simulates population structure, kinship, and family structure (Lu et al. 2020; Price et al. 2010). Our association result revealed a highly significant association for the two traits such as height and weight. Here, we presented the top five significant SNPs ($p < 5E-05$) associated with each trait (Table 1). The top lead significant variants are TIGRP2P201033_rs9187576 (height), BICF2P772349 (length), TIGRP2P119932_rs8658799 (LHR), BICF2P279062 (weight) which located in the intronic region of *HHIP*, the upstream region of *FGFR3*, intergenic region of *BEGAIN-DLK1*, and intronic region of *PTPN2* genes, respectively.

For the height trait, the *HHIP* gene (hedgehog interacting protein) involved in the development process has been previously reported to be associated with height and idiopathic short stature in children of the Korean population (Kim et al. 2010a). This gene has been detected for influencing bone and cartilage development, including skeletal development signaling pathway in human (Weedon and Frayling 2008; Zhao et al. 2010). Furthermore, *HHIP* is reported to play a role in lung branching morphogenesis during embryonic development (Chuang and McMahon 2003). Besides, our study enabled the identification of other noteworthy height associated candidate genes includes *LCORL* (ligand-dependent nuclear receptor corepressor like), *NCAPG* (non-smc condensin I complex subunit G) and *P4HA1* (prolyl 4-hydroxylase subunit alpha 1). The *LCORL* and *NCAPG* gene located on Chr3 correlated with height participates in spermatogenesis and cell cycle (Signer-Hasler et al. 2012; Takasuga 2016). These genes have constantly been recognized to influence human height in several large meta-analyses (Tetens et al. 2013). This *LCORL/NCAPG* locus has been previously identified for association with height and body-weight in horses and cattle (Signer-Hasler et al. 2012; Takasuga 2016). The *LCORL* locus, in particular, has previously been found to be associated with both height and body weight in dogs, and it has been estimated that this gene accounts for 15% of the bodyweight variation in the dog population (Plassais et al. 2019). Furthermore, the *LCORL* gene has been reported to be involved in different conformation trait-like head, frame, and neck development in horse as well as involved in trunk and hip development in human (Schröder 2010; Soranzo et al. 2009). The *P4HA1* gene located on Chr 4, which correlates with Jindo dog's height, is involved in collagen synthesis (Signer-Hasler et al. 2012; Vaysse et al. 2011). The *P4HA1* gene has not been reported in the dog before and the function of which in height regulation is also still unknown.

Bodyweight has been used as a surrogate for body size, which is the most striking trait for investigating genetic effects on quantitative traits. In human and even domestic animals, body size has been intensively studied (Allen et al. 2010; Boyko et al. 2010; Eigenmann et al. 1984; Hayward et al. 2016; Hoopes et al. 2012; Plassais et al. 2019; Rimbault et al. 2013; Sutter et al. 2007; Vaysse et al. 2011). Here, GWAS has helped us to identify candidate genes for body-weight in the Jindo breed. The protein tyrosine phosphatase non-receptor type 2 (*PTPN2*) gene located on Chr7 that significantly correlated with Jindo body-weight trait, is involved in the regulation of development of diabetic periodontitis (Hao et al. 2012). The result was in agreement with our previous GWAS study conducted on companion dog (Sheet et al. 2020). The *RASAL2* (RAS Protein Activator Like 2) gene on Chr7 that correlated with body-weight is closely associated with adipogenesis (Zhu et al. 2017). The *RASAL2* gene has been previously reported to be associated with increased body mass index in humans (Thorleifsson et al. 2009). Furthermore, we found that *PARPBP* and *IGF1* genes are strongly associated with body-weight in Jindo dog. The only gene, *IGF1* was close to two SNPs (BICF2S23655947 and BICF2S2336786) located in the 41134849–41239301 bp region of Chr15. The insulin like growth factor-1 (*IGF1*) gene is a well-known candidate gene associated with size variation, which has been reported in every dog body size GWAS as well as those of many other mammals and, of course, humans (Boyko et al. 2010; Eigenmann et al. 1984; Hayward et al. 2016; Hoopes et al. 2012; Plassais et al. 2019; Rimbault et al. 2013; Sutter et al. 2007; Vaysse et al. 2011).

IGF1 binds to a type 1 IGF receptor, a signal transducer for a tyrosine kinase (Sutter et al. 2007). This communication stimulates the growth of cells and causes cellular differentiation for the survival of organisms (Sutter et al. 2007). The Poly [ADP-ribose] polymerase 1 binding protein encoded gene, *PARPBP* has been reported before to be enriched in gene ontology terms, DNA repair and genome stability but the involvement in body-weight regulation is not yet confirmed (Engeland 2018). The small nuclear RNA, C/D Box 19 (*SNORD19*) have not reported previously to be associated with body weight was detected on Chr14 in multiple regions. The role of this gene in body-weight regulation is still not identified.

Unlike height and body weight, length and LHR have not been well studied and less is known about its underlying genetics. Our results report new as well as previously reported genes associated with length or LHR. We observed 5 and 13 significant correlated genetic variants for length and LHR, respectively. Among 5 loci (*FGFR3*, *PLCH1*, *IGF1*, *SCAND3*, and *ENSCAFG00000008038*), the *IGF1* gene on Chr15 was found to be common with body-size trait and had previously been reported in several studies (Boyko et al. 2010; Eigenmann et al. 1984; Hayward et al. 2016; Hoopes et al. 2012; Rimbault et al. 2013; Sutter et al. 2007; Vaysse et al. 2011). Previously, a genome-wide fine-mapping study within the Portuguese water dog population found marked evidence for a single *IGF1* SNP haplotype which is a major contributor to small body size in all small dogs (Sutter et al. 2007). The *IGF1* gene may substantially affect dog length by regulating the development and functioning of joint, and following controlling the body size (Frost et al. 2015; Sutter et al. 2007). The *FGFR3* gene encodes fibroblast growth factor receptor 3 located on Chr3 was the topmost associated candidate with length trait. Studies showed that fibroblast growth factor receptor is important for skeleton development (Su et al. 2014). Mutations in this gene cause human skeleton dysplasia, including developmental delay and achondroplasia (Martinez et al. 2000; Su et al. 2014). Earlier, a selective sweep mapping study on detection of the genetic interval containing putative genes causing foreshortened limbs in Dachshunds dog by Pollinger et al., found a large sweep region in the vicinity of the *FGRF3* gene, proposing that the causative mutation in dogs is in a gene or regulatory region closely associated to *FGRF3* (Pollinger et al. 2005). This condition can lead to short stature, macrocephaly, particularly in the proximal lower and upper limbs (Su et al. 2014). Among the 13 LHR associated genome-wide significant genes, the *EFEMP1* gene (EGF containing fibulin extracellular matrix protein 1) was showed to be associated with human height (Liu et al. 2010). We identified another significant gene, *DLK1* on Chr8 which is one the most strongly associated loci with LHR. The *DLK1* (delta-like non-canonical notch ligand 1) encodes a growth factor containing transmembrane protein which works as a controller of cell growth (Traustadóttir et al. 2019). Furthermore, this gene is largely known for its contribution to adipogenesis and organism development (Traustadóttir et al. 2019).

Several of the genes that were observed to be associated with substantially linked SNPs in our study were enriched in GO terms, e.g. thyroid hormone synthesis, extracellular matrix binding, actin cytoskeleton ($p < 0.05$). Our result exposed that the enrichment of gene-sets which worked together in a network to accomplish specific molecular processes (Sheet et al. 2020). Among 9 GO terms, the nucleus was the enriched cellular component with the highest number of genes (Supplementary Table S2). It was previously found to be associated with weight trait in companion dogs (Sheet et al. 2020). The actin cytoskeleton is an essential cellular component for muscle development as it transforms myotubes into myofibers (Lu et al. 2020; Picard et al. 2010). Muscle development is a complex process and it mainly depends on the hypertrophy and proliferation of muscle fibers (Picard et al. 2010). The actin and actin-binding protein with another component of the cytoskeleton including the microtubule works together in a network to help essential cellular processes like axonal growth and cell migration (Picard et al. 2010). The extracellular matrix has been associated with body-weight regulation as it helps cells in proliferation (Roumans et al. 2015). Thyroid hormone synthesis is closely correlated with growth and development (Tarim 2011). It was noticed that the overlap in GO terms is largely housekeeping pathways, except for thyroid hormone synthesis. The genes enriched in thyroid hormone synthesis were *BDNF* (brain-derived neurotrophic factor), *CBLN1* (cerebellin 1 precursor), *SFRP1* (secreted frizzled-related protein 1), *IGF1* (insulin-like growth factor-1), *NTRK2* (neurotrophic receptor tyrosine kinase 2), *ACACA* (acetyl-coa carboxylase alpha), *FGFR3* (fibroblast growth factor receptor 3). Polymorphisms in two of these genes (*IGF1* and *FGFR3*) were found to be associated with body size and leg length in dogs, respectively (Hayward et al. 2016; Pollinger et al. 2005; Vaysse et al. 2011). Boule et al. described that the glucose homeostasis has potential role in weight regulation and it can predicts the weight gain (Boulé et al. 2008).

The KEGG pathway analysis always has been used to access the group of genes associated with a particular trait. This method can identify enrichments by pooling data from multiple genetic SNPs, particularly through individual genetic SNP that do not meet a specific significance threshold (Zhu and Stephens 2018). Our KEGG pathway analysis has successfully detected several significantly enriched pathways. Some shared pathways between traits were noticed. It has been proposed that the regulation of correlated traits is more likely controlled by similar or related pathways (Pei et al. 2019). Here, we identified seven overlapping pathways enriched to all traits (height, length, LHR, and body-weight). Interestingly, six of those pathways were hormone-related pathways, such as the insulin

secretion pathway, the estrogen signaling pathway, the aldosterone synthesis and secretion, the glucagon signaling pathway, the insulin resistance pathway, and the thyroid hormone synthesis pathway. The hormone is a key regulatory system for growth and development. Insulin and thyroid hormones are the most vital endogenous regulators of growth and skeletal development after growth hormone. The significance of the thyroid hormone for skeletal development and maintenance has been well documented in several studies (Tarim 2011). The previous study showed that thyroid hormones control the metabolism of body proteins along with other hormones like growth hormone, insulin-like growth factor-1 (IGF-1), insulin, and glucocorticoids, and thus regulate the growth and development process (Tarim 2011). Insulin is a peptide hormone that has several effects on the metabolism of fats, proteins, and carbohydrates (Do et al. 2017). Growth hormone receptor (GHR) sensitivity and the level of insulin-like growth factor 1 (IGF-1) can also be affected by insulin, consequently influencing the level of growth hormone (Qiu et al. 2017). The insulin resistance pathway has been previously reported to be associated with height and body-weight trait in human (Deng et al. 2012). The glucagon signaling pathway is a pathway that strongly affects the body-weight trait by increasing blood glucose level in the blood through glycogenolysis and gluconeogenesis (Jiang and Zhang 2003). In addition, we have identified another signaling pathway shared between traits is cGMP-PKG signaling pathway. The cGMP-PKG signaling pathway reported to be associated with contraction of vascular smooth muscle cells. It has previously been found in our previous study to be correlated with body-weight trait in dog (Sheet et al. 2020).

Conclusion

In this study, we have discovered several new unreported as well as previously identified markers, genes and pathways that might influence the variation in height, length, LHR, and body-weight of Jindo dog using GWAS with pathway-based approaches. The significantly associated markers harboring *HHIP*, *LCORL*, *NCAPG*, *IGFI*, *FGFR3*, *DLK1*, *EFEMP1*, *PTPN2*, *IGFI*, and *RASAL2* genes are probably the most attractive candidate genes as their functions are strongly related for height, length, and body-weight phenotypes. Interestingly, the identified overlapping enriched pathways were found to be mainly hormone-related pathways (insulin secretion pathway, the estrogen signaling pathway, the aldosterone synthesis and secretion, the glucagon signaling pathway, insulin resistance pathway, and the thyroid hormone synthesis pathway) which participate in growth and development process. Integrating gene expression data along with our association data used in this study may expose the true genetic mechanisms underlying height, length, and body-weight variation in Jindo dog.

Declarations

Authors' contributions SS and BHC: Conceptualization and carried out research, JSK, MJK, NYK, YJL, MRP, SJL, and JMK: carried out research, resources, and software, SO and BHC: data curation, SS and BHC: analyzed data and wrote the manuscript. All the authors read and approved the final manuscript.

Funding This research was supported by the funds from the AGENDA project (grant no. PJ013415012020) of the National Institute of Animal Science, Rural Development Administration, Republic of Korea.

Data availability Our manuscript has data included as electronic supplementary material. Data will be sent to the interested researcher upon request.

Code availability Code will be sent to the interested researcher upon request.

Conflicts of Interest The authors declare that they have no conflict of interest.

Ethical approval Blood samples were collected from them by following relevant guidelines formulated by the Institutional Animal Care and Use Committee (IACUC) of the National Institute of Animal Science (NIAS, RDA, Wanju-gun, South Korea), and protocol consent was obtained for 'Development of early diagnosis technology for degenerative muscular skeleton system in special-purpose dog' project.

Consent to participate Not applicable

Consent for publication Not applicable

Acknowledgements This study was supported by 2020 the RDA Fellowship Program of the National Institute of Animal Science, Rural Development Administration, Republic of Korea. All Jindo dog blood samples were supplied from the Korean Jindo-gun, Jeollanam-do, Republic of Korea.

References

1. Albarella U (2003) Dogs in Antiquity. Anubis to Cerberus: the origins of the domestic dog. Brewer D, Clark T, Phillips A. Aris & Phillips: Warminster, 2001. ISBN 0-85668-704-9 vol 13. John Wiley & Sons, Ltd. Chichester, UK. <https://doi.org/10.1002/oa.649>
2. Allen HL et al (2010) Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* 467:832–838
3. Bannasch DL, Baes CF, Leeb T (2020) Genetic Variants Affecting Skeletal Morphology in Domestic Dogs. *Trends Genet*
4. Boulé NG, Chaput JP, Doucet E, Richard D, Després JP, Bouchard C, Tremblay A (2008) Glucose homeostasis predicts weight gain: prospective and clinical evidence. *Diabetes Metab* 24:123–129
5. Boyko AR et al (2010) A simple genetic architecture underlies morphological variation in dogs. *Plos Biol* 8:e1000451
6. Bray EE, Otto CM, Udell MA, Hall NJ, Johnston AM, MacLean EL (2021) Enhancing the Selection and Performance of Working Dogs. *Front Vet Sci* 8:430
7. Chan Y et al (2015) Genome-wide analysis of body proportion classifies height-associated variants by mechanism of action and implicates genes important for skeletal development. *Am J Hum Genet* 96:695–708
8. Chase K, Carrier DR, Adler FR, Jarvik T, Ostrander EA, Lorentzen TD, Lark KG (2002) Genetic basis for systems of skeletal quantitative traits: principal component analysis of the canid skeleton. *Proc Natl Acad Sci* 99:9930–9935
9. Chuang P-T, McMahon AP (2003) Feedback control of mammalian Hedgehog signaling by the Hedgehog-binding protein, *Hip1*, modulates *Fgf* signaling during branching morphogenesis of the lung. *Genes Dev* 17:342–347
10. Club AK (1998) The complete dog book. Howell Books, New York
11. Dadousis C, Pegolo S, Rosa G, Gianola D, Bittante G, Cecchinato A (2017) Pathway-based genome-wide association analysis of milk coagulation properties, curd firmness, cheese yield, and curd nutrient recovery in dairy cattle. *J Dairy Sci* 100:1223–1231
12. Deng H-Z, Deng H, Su Z, Li Y-H, Ma H-M, Chen H-S, Du M-L (2012) Insulin resistance and adiponectin levels are associated with height catch-up growth in pre-pubertal Chinese individuals born small for gestational age. *Nutr Metab* 9:1–7
13. Do D, Bissonnette N, Lacasse P, Miglior F, Sargolzaei M, Zhao X, Ibeagha-Awemu E (2017) Genome-wide association analysis and pathways enrichment for lactation persistency in Canadian Holstein cattle. *J Dairy Sci* 100:1955–1970
14. Drögemüller M et al (2014) A mutation in the *FAM83G* gene in dogs with hereditary footpad hyperkeratosis (HFH). *PLoS Genet* 10:e1004370
15. Eigenmann J, Patterson D, Froesch E (1984) Body size parallels insulin-like growth factor I levels but not growth hormone secretory capacity. *Eur J Endocrinol* 106:448–453
16. Engeland K (2018) Cell cycle arrest through indirect transcriptional repression by p53: I have a DREAM. *Cell Death Differ* 25:114–132
17. Frost A, Proskura W, Barnik J, Szatkowska I, Dybus A (2015) Variability within IGF1 and its linkage to body size and predisposition to hip dysplasia in dogs. *Indian J Anim Sci* 85:411–413
18. Hao L, Wei L, Yi D, Bin G, Honghui X, Min F, Qi W (2012) Affect of protein tyrosine phosphatase non-receptor type 2 and nuclear factor- κ B on periodontal destruction with diabetes West China. *Czas Stomatol* 30
19. Hayward JJ et al (2016) Complex disease and phenotype mapping in the domestic dog. *Nature Commun* 7:1–11
20. Helmink S, Rodriguez-Zas S, Shanks R, Leighton E (2001) Estimated genetic parameters for growth traits of German shepherd dog and Labrador retriever dog guides. *J Ani Sci* 79:1450–1456
21. Hoopes BC, Rimbault M, Liebers D, Ostrander EA, Sutter NB (2012) The insulin-like growth factor 1 receptor (*IGF1R*) contributes to reduced size in dogs. *Mamm genome* 23:780–790
22. Jiang G, Zhang BB (2003) Glucagon and regulation of glucose metabolism. *Am J Physiol Endocrinol Metab* 284:E671–E678
23. Jones P, Chase K, Martin A, Davern P, Ostrander EA, Lark KG (2008) Single-nucleotide-polymorphism-based association mapping of dog stereotypes. *Genetics* 179:1033–1044
24. Kim J-J et al (2010a) Identification of 15 loci influencing height in a Korean population. *J Hum Genet* 55:27–31
25. Kim RN et al (2012) Genome analysis of the domestic dog (Korean Jindo) by massively parallel sequencing. *DNA Res* 19:275–288
26. Kim Y-M, Am AE-A, Hwang S-H, Lee J-H, Lee S-M (2009) Risk factors of relinquishment regarding canine behavior problems in South Korea. *Berl tierarztl Wschr* 122:1–7

27. Kim Y-M, Kim S-A, Lee S-M, Choi Y-J, Kim B-J, Shin N-S (2010b) Canine behavioral problems and their effect on relinquishment of the Jindo dog. *J Vet Sci* 11:345
28. Lazarowski L et al (2018) Investigation of the behavioral characteristics of dogs purpose-bred and prepared to perform Vapor Wake® detection of person-borne explosives. *Front Vet Sci* 5:50
29. Lee C-M, Song D-W, Ro W-B, Kang M-H, Park H-M (2019a) Genome-wide association study of degenerative mitral valve disease in Maltese dogs. *J Vet Sci* 20:63–71
30. Lee E et al (2019b) Genome-wide enriched pathway analysis of acute post-radiotherapy pain in breast cancer patients: a prospective cohort study. *Hum Genomics* 13:28
31. Liu JZ et al (2010) Genome-wide association study of height and body mass index in Australian twin families. *Twin Res Hum Genet* 13:179–193
32. Lu Z et al (2020) Genome-wide association study of body-weight traits in chinese fine-wool sheep. *Animals* 10:170
33. Martinez S, Valdes J, Alonso RA (2000) Achondroplastic dog breeds have no mutations in the transmembrane domain of the *FGFR-3* gene. *Can J Vet Res* 64:243
34. McGreevy PD, Georgevsky D, Carrasco J, Valenzuela M, Duffy DL, Serpell JA (2013) Dog behavior co-varies with height, bodyweight and skull shape. *PLoS one* 8:e80529
35. Momozawa Y et al (2020) Genome wide association study of 40 clinical measurements in eight dog breeds. *Sci Rep* 10:1–11
36. Parker HG et al (2009) An expressed *fgf4* retrogene is associated with breed-defining chondrodysplasia in domestic dogs. *Sci* 325:995–998
37. Pei G, Sun H, Dai Y, Liu X, Zhao Z, Jia P (2019) Investigation of multi-trait associations using pathway-based analysis of GWAS summary statistics. *BMC Genom* 20:43–54
38. Picard B, Berri C, Lefaucheur L, Molette C, Sayd T, Terlouw C (2010) Skeletal muscle proteomics in livestock production. *Brief Funct Genomics* 9:259–278
39. Pilot M et al (2015) On the origin of mongrels: evolutionary history of free-breeding dogs in Eurasia. *Proc R Soc Lond [Biol]* 282:20152189
40. Plassais J et al (2019) Whole genome sequencing of canids reveals genomic regions under selection and variants influencing morphology. *Nat Commun* 10:1–14
41. Plassais J, Rimbault M, Williams FJ, Davis BW, Schoenebeck JJ, Ostrander EA (2017) Analysis of large versus small dogs reveals three genes on the canine X chromosome associated with body-weight, muscling and back fat thickness. *PLoS Genet* 13:e1006661
42. Pollinger JP, Bustamante CD, Fledel-Alon A, Schmutz S, Gray MM, Wayne RK (2005) Selective sweep mapping of genes with large phenotypic effects. *Genome Res* 15:1809–1819
43. Price AL, Zaitlen NA, Reich D, Patterson N (2010) New approaches to population stratification in genome-wide association studies. *Nat Rev Genet* 11:459–463
44. Qiu H, Yang J-K, Chen C (2017) Influence of insulin on growth hormone secretion, level and growth hormone signalling. *Sheng Li Xue Bao* 69:541–556
45. Quignon P et al (2009) Fine mapping a locus controlling leg morphology in the domestic dog. In: *Cold Spring Harbor symposia on quantitative biology*, vol 74. Cold Spring Harbor Laboratory Press, New York, pp 327–333. <https://doi.org/10.1101/sqb.2009.74.009>
46. Rimbault M et al (2013) Derived variants at six genes explain nearly half of size reduction in dog breeds. *Genome Res* 23:1985–1995
47. Roumans NJ et al (2015) Variation in extracellular matrix genes is associated with weight regain after weight loss in a sex-specific manner. *Genes Nutr* 10:1–9
48. Schmidt S, Hartung P, Rath J, Piepho J H-P (2019) Estimating broad-sense heritability with unbalanced data from agricultural cultivar trials. *Crop Sci* 59:525–536
49. Schröder W (2010) Athletic performance and conformation in Hanoverian warmblood horses-population genetic and genome-wide association analyses. *Bibliothek der Tierärztlichen Hochschule Hannover*
50. Sheet S, Krishnamoorthy S, Cha J, Choi S, Choi B-H (2020) Identification of Candidate Genes and Pathways Associated with Obesity-Related Traits in Canines via Gene-Set Enrichment and Pathway-Based GWAS Analysis. *Animals* 10:2071

51. Signer-Hasler H, Flury C, Haase B, Burger D, Simianer H, Leeb T, Rieder S (2012) A genome-wide association study reveals loci influencing height and other conformation traits in horses. *PLoS one* 7:e37282
52. Soranzo N et al (2009) Meta-analysis of genome-wide scans for human adult stature identifies novel Loci and associations with measures of skeletal frame size. *PLoS Genet* 5:e1000445
53. Spady TC, Ostrander EA (2008) Canine behavioral genetics: pointing out the phenotypes and herding up the genes. *Am J Hum Genet* 82:10–18
54. Srikanth K et al (2020) A Gene-Set Enrichment and Protein–Protein Interaction Network-Based GWAS with Regulatory SNPs Identifies Candidate Genes and Pathways Associated with Carcass Traits in Hanwoo Cattle. *Genes* 11:316
55. Stockard CR (1941) The genetic and endocrine basis for differences in form and behaviour as elucidated by studies of contrasted pure-line dog breeds and their hybrids. With special contributions on behaviour by OD Anderson and WT James The genetic and endocrine basis for differences in form and behaviour as elucidated by studies of contrasted pure-line dog breeds and their hybrids With special contributions on behaviour by OD Anderson and WT James. Wistar Institute of Anatomy and Biology, Philadelphia. <https://www.cabdirect.org/cabdirect/abstract/19430100872>
56. Stone HR, McGreevy PD, Starling MJ, Forkman B (2016) Associations between domestic-dog morphology and behaviour scores in the dog mentality assessment. *PLoS one* 11:e0149403
57. Su N, Jin M, Chen L (2014) Role of *FGF/FGFR* signaling in skeletal development and homeostasis: learning from mouse models. *Bone Res* 2:1–24
58. Sutter NB et al (2007) A single *IGF1* allele is a major determinant of small size in dogs. *Science* 316:112–115
59. Takasuga A (2016) *PLAG1* and *NCAPG-LCORL* in livestock. *Ani Sci J* 87:159–167
60. Tarım Ö (2011) Thyroid hormones and growth in health and disease. *J Clin Res Pediatr Endocrinol* 3:51
61. Tetens J, Widmann P, Kühn C, Thaller G (2013) A genome-wide association study indicates *LCORL/NCAPG* as a candidate locus for withers height in German Warmblood horses. *Ani Genet* 44:467–471
62. Thorleifsson G et al (2009) Genome-wide association yields new sequence variants at seven loci that associate with measures of obesity. *Nature Genet* 41:18
63. Traustadóttir G, Lagoni LV, Ankerstjerne LBS, Bisgaard HC, Jensen CH, Andersen DC (2019) The imprinted gene Delta like non-canonical Notch ligand 1 (*Dlk1*) is conserved in mammals, and serves a growth modulatory role during tissue development and regeneration through Notch dependent and independent mechanisms. *Cytokine Growth Factor Rev* 46:17–27
64. Vaysse A et al (2011) Identification of genomic regions associated with phenotypic variation between dog breeds using selection mapping. *PLoS Genet* 7:e1002316
65. Weedon MN, Frayling TM (2008) Reaching new heights: insights into the genetics of human stature. *Trends Genet* 24:595–603
66. Young A, Bannasch D (2006) Morphological variation in the dog. *Cold spring harbor monograph series* 44:47. doi:<http://dx.doi.org/10.1101/0.47-65>
67. Zhao J et al (2010) The role of height-associated loci identified in genome wide association studies in the determination of pediatric stature. *BMC Med Genet* 11:1–7
68. Zhu X, Stephens M (2018) Large-scale genome-wide enrichment analyses identify new trait-associated genes and pathways across 31 human phenotypes. *Nat Commun* 9:1–1
69. Zhu X, Xie S, Xu T, Wu X, Han M (2017) *Rasal2* deficiency reduces adipogenesis and occurrence of obesity-related disorders. *Mol Metab* 6:494–502

Figures

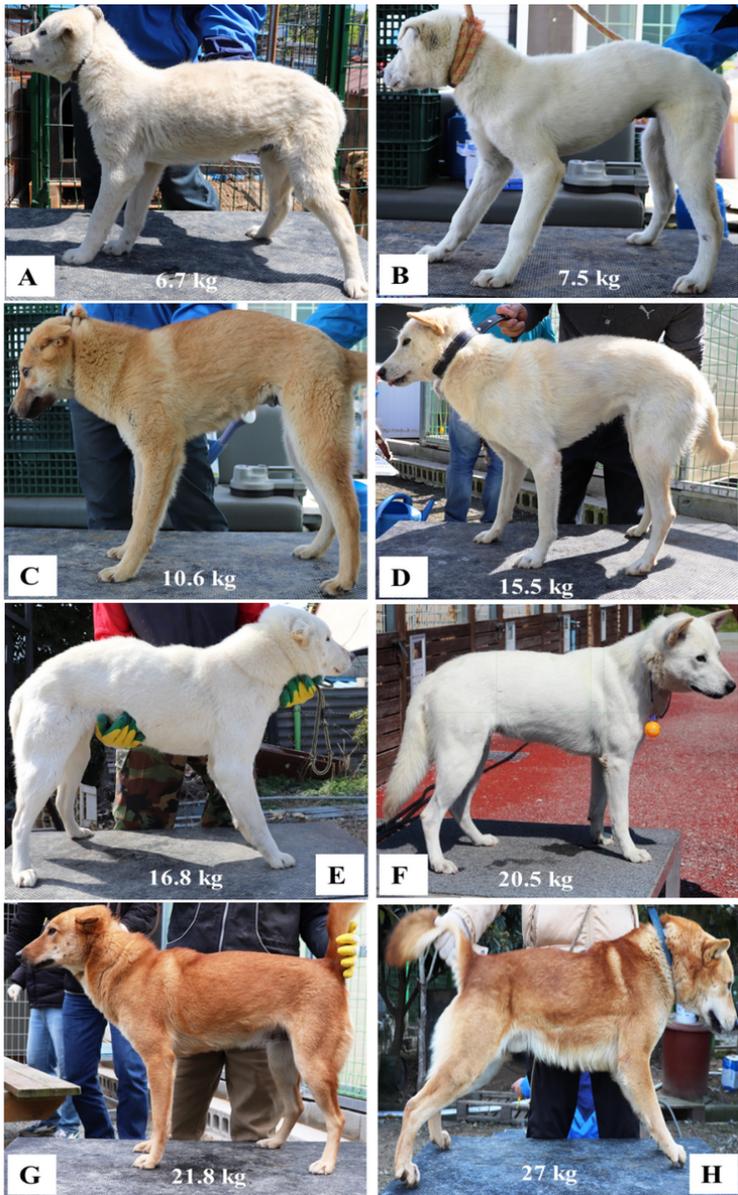


Figure 1

The phenotypic variation of Jindo dog. (A). Male Jindo dog with height 35.5 cm, length 38.5 cm, LHR 114.1, (B). Female Jindo dog with height 35 cm, length 40 cm, LHR 114.3, (C). Female Jindo dog with height 42 cm, length 43 cm, LHR 102.4, (D). Female Jindo dog with height 47 cm, length 48 cm, LHR 102.1, (E). Female Jindo dog with height 50 cm, length 54 cm, LHR 108, (F). Female Jindo dog with height 51.5 cm, length 54.4 cm, LHR 105.6, (G). Female Jindo dog with height 49 cm, length 52 cm, LHR 106.1, (H). Male Jindo dog with height 52.0 cm, length 55.5 cm, LHR 106.7.

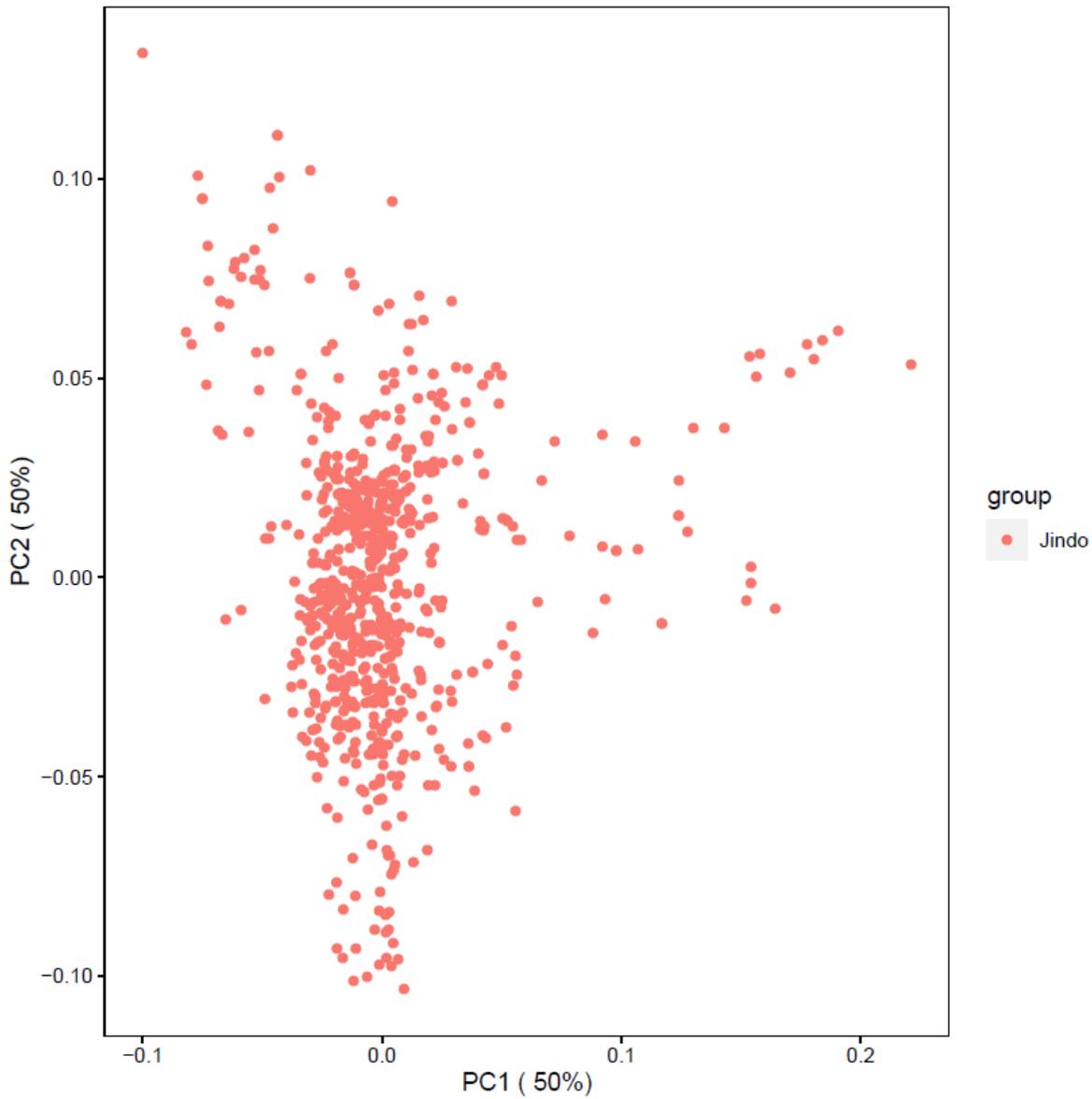


Figure 2

Population structure for all Jindo dog.

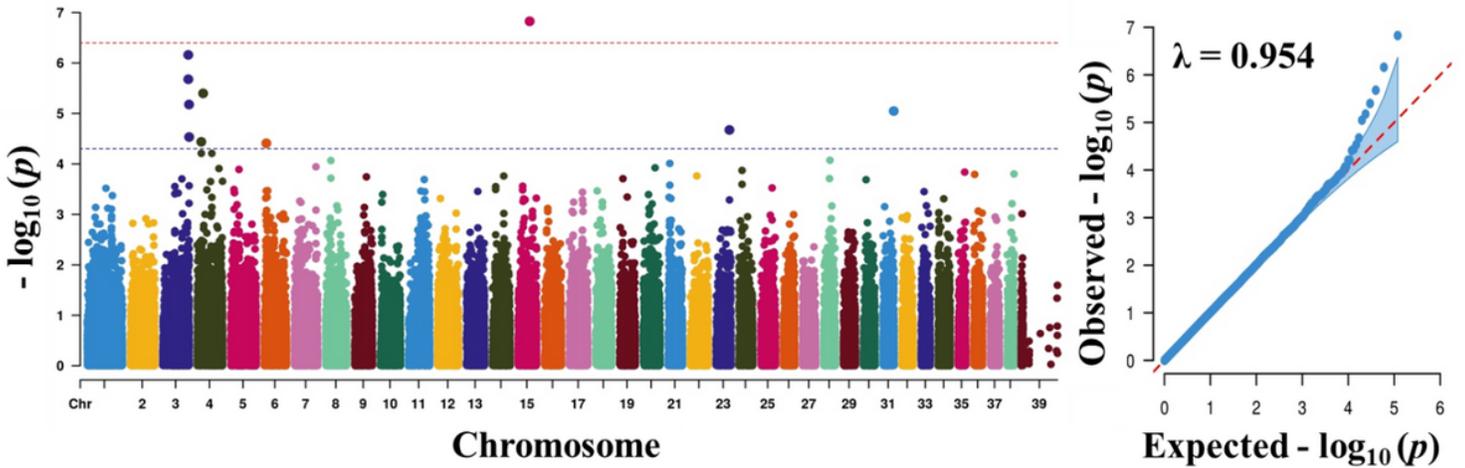


Figure 3

Manhattan plots of mapped single nucleotide polymorphism (SNP) markers associating with height. Red line designates the Bonferroni corrected significance threshold level of $p < 4E - 07$ and blue line shows genome wide significant threshold level of $p < 5E - 05$. The quantile-quantile (Q-Q) plot of the GWAS shows on the right side. GWAS, genome-wide association study.

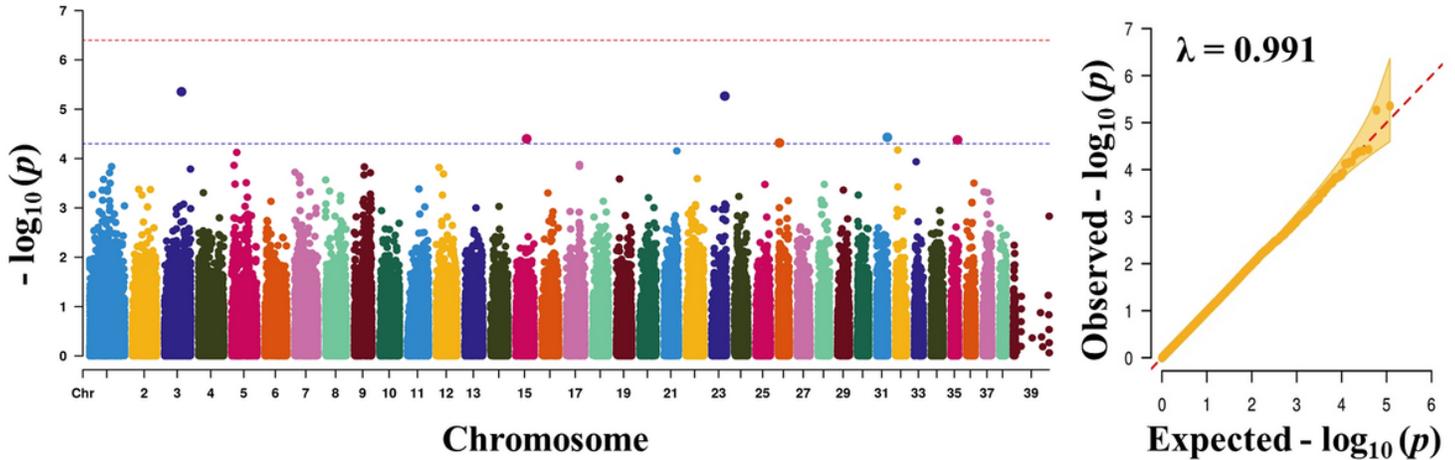


Figure 4

Manhattan plots of mapped single nucleotide polymorphism (SNP) markers associating with length. Red line designates the Bonferroni corrected significance threshold level of $p < 4E - 07$ and blue line shows genome wide significant threshold level of $p < 5E - 05$. The quantile-quantile (Q-Q) plot of the GWAS shows on the right side. GWAS, genome-wide association study.

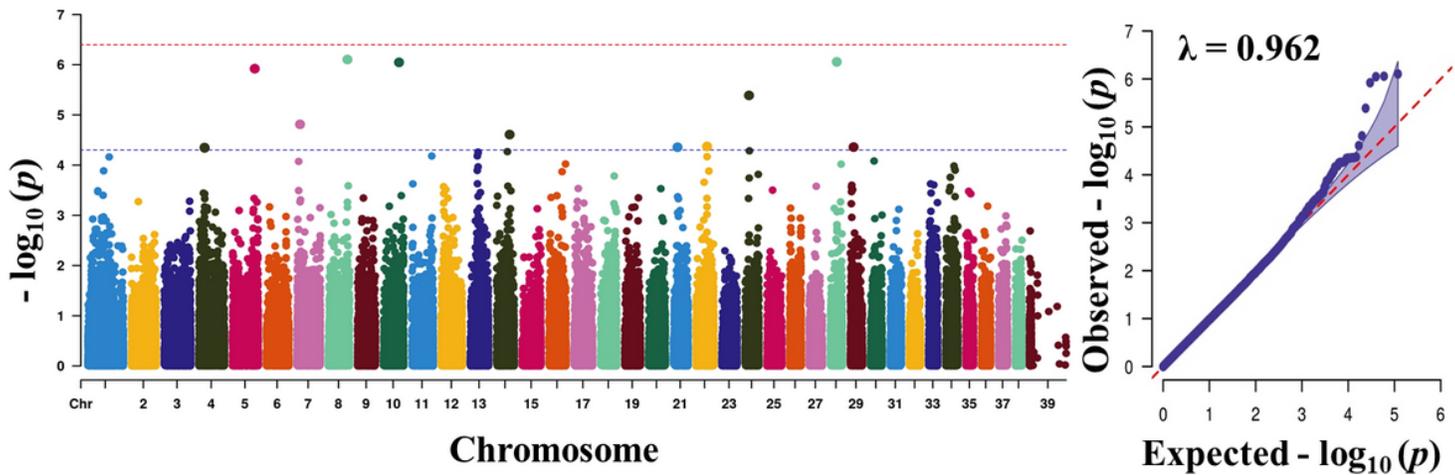


Figure 5

Manhattan plots of mapped single nucleotide polymorphism (SNP) markers associating with LHR. Red line designates the Bonferroni corrected significance threshold level of $p < 4E - 07$ and blue line shows genome wide significant threshold level of $p < 5E - 05$. The quantile-quantile (Q-Q) plot of the GWAS shows on the right side. GWAS, genome-wide association study. LHR, length to height ratio.

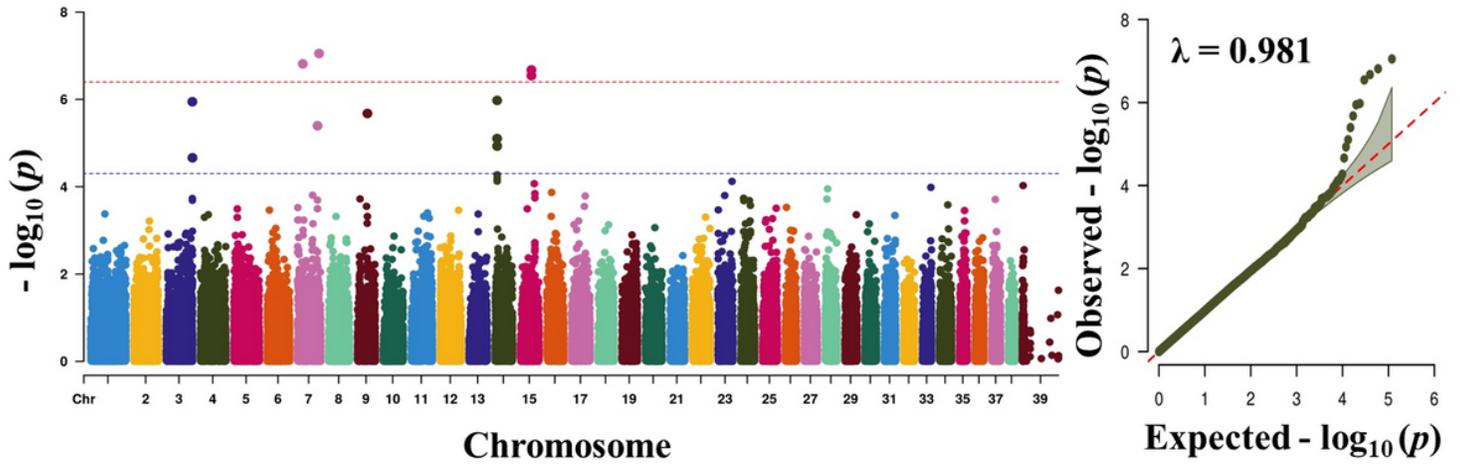


Figure 6

Manhattan plots of mapped single nucleotide polymorphism (SNP) markers associating with body-weight. Red line designates the Bonferroni corrected significance threshold level of $p < 4E - 07$ and blue line shows genome wide significant threshold level of $p < 5E - 05$. The quantile-quantile (Q-Q) plot of the GWAS shows on the right side. GWAS, genome-wide association study.

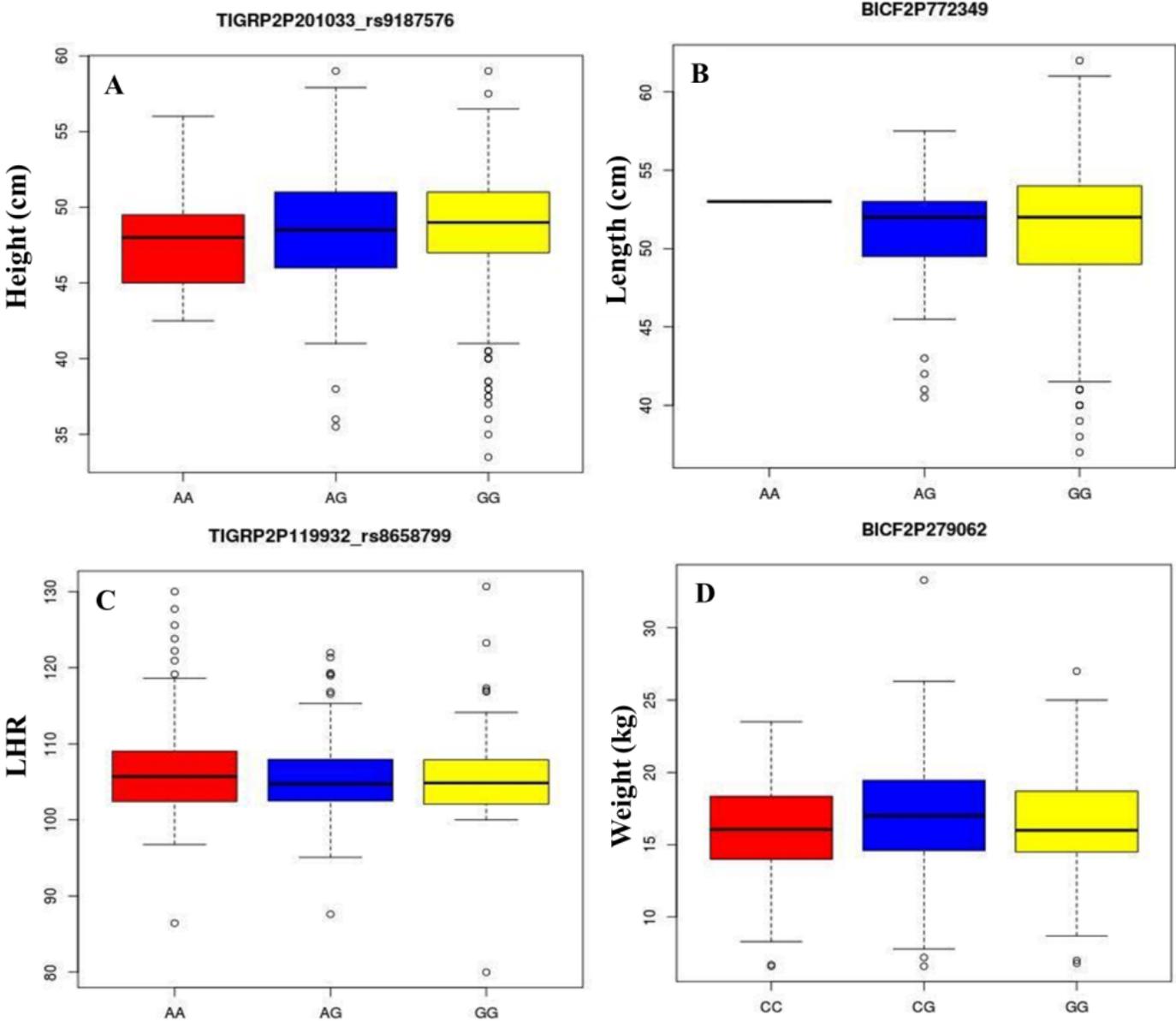


Figure 7

Genotype-phenotype correlation. The box plots show the distribution of phenotypes among the different genotypes of top variants associated to height, length, length to height ratio, and body-weight.

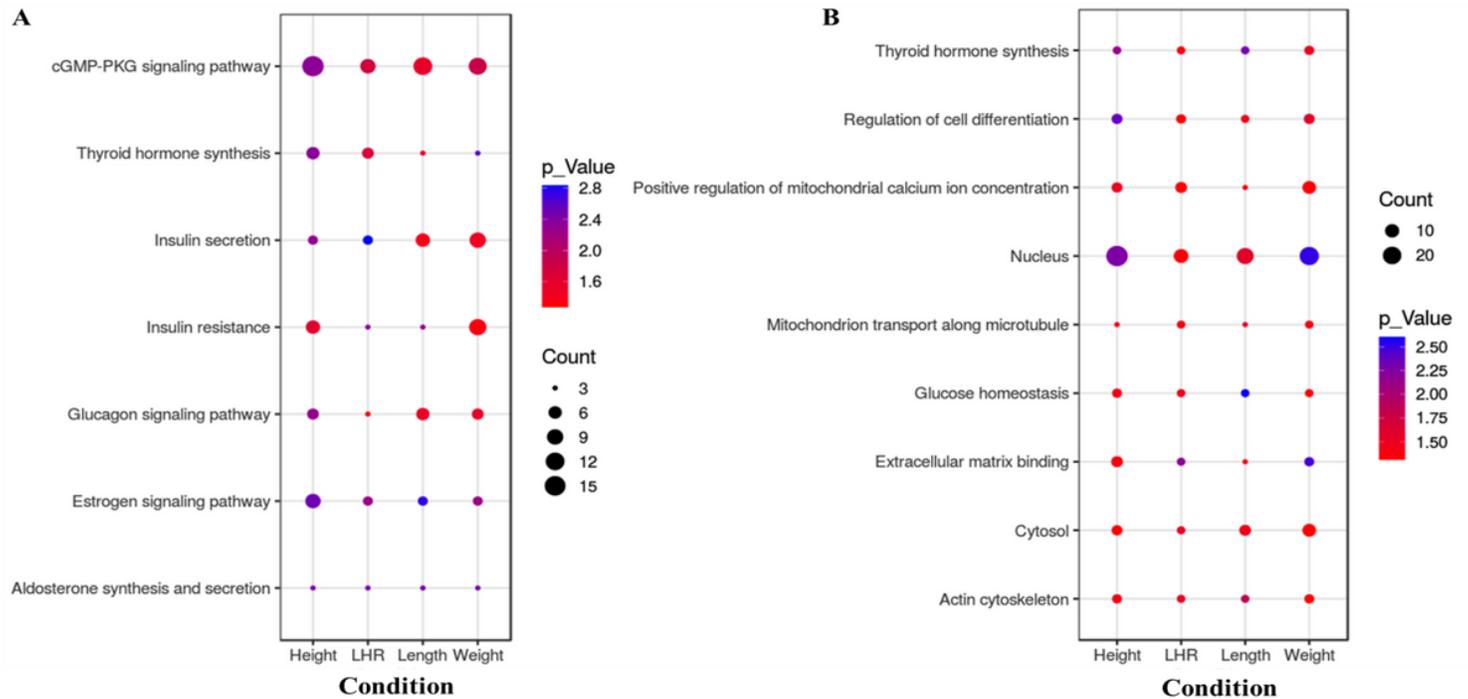


Figure 8

The significantly ($p < 0.05$) enriched 7 KEGG pathways (A) and 9 GO terms (B) shared among traits.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryTableS1.xlsx](#)
- [SupplementaryTableS2.xlsx](#)