

RESEARCH

Open Access



# Genome-wide association study of 17 serum biochemical indicators in a chicken F<sub>2</sub> resource population

Haijie Song<sup>1†</sup>, Wenting Li<sup>1,2†</sup>, Yuanfang Li<sup>1</sup>, Bin Zhai<sup>1</sup>, Yujie Guo<sup>1</sup>, Yi Chen<sup>1</sup>, Ruili Han<sup>1,2</sup>, Guirong Sun<sup>1,2</sup>, Ruirui Jiang<sup>1,2</sup>, Zhuanjian Li<sup>1,2</sup>, Fengbin Yan<sup>1,2</sup>, Guoxi Li<sup>1,2</sup>, Xiaojun Liu<sup>1,2</sup>, Yanhua Zhang<sup>1,2\*</sup>, Yadong Tian<sup>1,2\*</sup> and Xiangtao Kang<sup>1,2\*</sup>

## Abstract

**Background** Serum biochemical indicators are often regarded as direct reflections of animal metabolism and health. The molecular mechanisms underlying serum biochemical indicators metabolism of chicken (*Gallus Gallus*) have not been elucidated. Herein, we performed a genome-wide association study (GWAS) to identify the variation associated with serum biochemical indicators. The aim of this research was to broaden the understanding of the serum biochemical indicators in chickens.

**Results** A GWAS of serum biochemical indicators was carried out on 734 samples from an F<sub>2</sub> Gushix Anka chicken population. All chickens were genotyped by sequencing, 734 chickens and 321,314 variants were obtained after quality control. Based on these variants, a total of 236 single-nucleotide polymorphisms (SNPs) on 9 chicken chromosomes (GGAs) were identified to be significantly ( $-\log_{10}(P) > 5.72$ ) associated with eight of seventeen serum biochemical indicators. Ten novel quantitative trait loci (QTLs) were identified for the 8 serum biochemical indicator traits of the F<sub>2</sub> population. Literature mining revealed that the *ALPL*, *BCHE*, *GGT2/GGT5* genes at loci GGA24, GGA9 and GGA15 might affect the alkaline phosphatase (AKP), cholinesterase (CHE) and  $\gamma$ -glutamyl transpeptidase (GGT) traits, respectively.

**Conclusion** The findings of the present study may contribute to a better understanding of the molecular mechanisms of chicken serum biochemical indicator regulation and provide a theoretical basis for chicken breeding programs.

**Keywords** Chicken, Genome-wide association study, 17 serum biochemical indicators

<sup>†</sup>These authors contributed equally to this work.

\*Correspondence:  
Yanhua Zhang  
308762567@qq.com  
Yadong Tian  
ydtian111@163.com

Xiangtao Kang  
xtkang2001@263.net

<sup>1</sup>College of Animal Science and Technology, Henan Agricultural University, No.15 Longzihu University Area, Zhengzhou New District, 450002 Zhengzhou, China

<sup>2</sup>Henan Key Laboratory for Innovation and Utilization of Chicken Germplasm Resources, Henan Agricultural University, No.15 Longzihu University Area, Zhengzhou New District, 450002 Zhengzhou, China



## Background

Conventional breeding methods are based mainly on phenotypic selection and have certain limitations. Fat composition, immunity and some other traits are difficult to measure directly and can be measured only in siblings or by slaughter, resulting in problems such as long generation intervals and high costs. In the chicken breeding industry, development of convenient measurement indicators that reflect physiological fat metabolism, immune function and other parameters is necessary. Serum biochemical indicators are considered reflections of animal metabolism and health status and have great significance in livestock and poultry production research. These indicators are also easy to measure. Analyses of serum biochemical indicators are helpful for indirect selection of some traits, assessment of animal health, early diagnosis of animal diseases and other procedures and effectively compensate for the shortcomings of long-term and high-cost conventional breeding. In addition, molecular breeding can accelerate the speed of animal breeding. Therefore, screening molecular markers and analyzing the genetic basis of serum biochemical indicators through GWAS is a promising avenue for exploiting genomic information for commercial breeding of chickens.

Serum enzymes are responsible for catalyzing various physiological and biochemical processes in animals, and their levels and activities reflect growth, development and metabolism levels to a certain extent. Alkaline phosphatase (AKP) is a metalloenzyme superfamily member that hydrolyzes or transphosphorylates various biomolecules and exhibits its highest biological activity in suitable alkaline environments [1]. AKP is widespread in nature and exists in almost all animal tissues. The levels of AKP are higher in the skeleton, liver and bile duct than in other tissues. AKP in serum mainly originates in the skeleton and liver. AKP levels are often used to diagnose and monitor skeletal or liver diseases in the clinic [2]. Valgus-varus deformity (VVD) is a leg disease which lateral or middle deviation of the tibiotarsus or tarso-metatarsus, accompanied by abnormal changes in serum AKP levels. Case-control (VVD broilers-sound broilers)-based GWAS identified 43 SNPs located on Chr24 (0.22 Mb –1.79 Mb) that were significantly associated with AKP [3]. In the chicken F<sub>2</sub> resource population, the direct relationships between AKP levels and candidate genes and corresponding QTLs have not been reported at present. Cholinesterase (CHE) is an esterase that can break down choline esters and act as a neurotransmitter. Kidney disease, obesity, fatty liver, hyperthyroidism, and other conditions can lead to increased CHE levels. Measurement of serum CHE activity [4], serum alanine transaminase (ALT) levels and serum aspartate aminotransferase (AST) levels [5] is important for diagnosis of

liver diseases, and ALT is a sensitive biochemical marker used to evaluate damage to hepatic parenchymal cells [5].

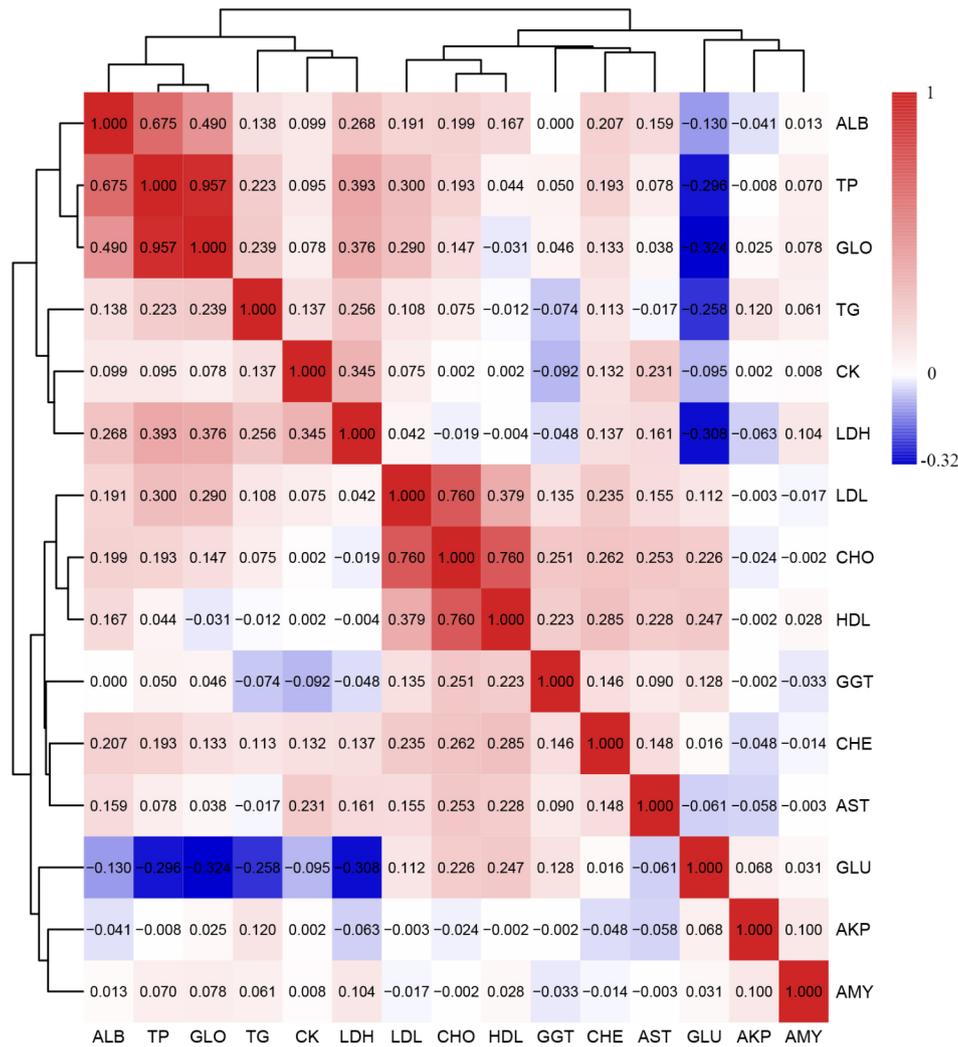
Blood lipid levels reflect overall lipid metabolism and health status. Abnormal blood lipid levels are associated with cardiovascular disease, familial hypercholesterolemia and diabetes. Clinical tests for total cholesterol (CHO), low-density lipoprotein cholesterol (LDL), high density lipoprotein cholesterol (HDL) and triglycerides (TGs) are widely used for cardiovascular disease risk assessment [6, 7]. Another important role of serum biochemical indicators is use them as indirect indicators for certain economic traits in animal breeding, such as the use of blood LDL as selection indicators for reducing abdominal fat content in broilers [8]. Fatty pigs have higher serum total protein (TP) levels than lean pigs, and serum TP level has been shown to be the best marker for early fat assessment in pigs [9, 10]. Multiple serum biochemical indicators have been identified as biomarkers for predicting important economic traits or general health in animals and humans, and deciphering the genetic basis of these biomarkers using GWAS is important for potential applications in animal breeding. Several genes associated with the electrolyte levels in serum were identified by case-control (health-disease) -based GWAS [11, 12]. Several early studies have reported associated QTLs for blood components in chickens [13–15]. As far as we know, no published study has systematically demonstrated the associations between the 17 serum biochemical indicators and genomic loci using a GWAS of SNPs in chickens.

Biochemical indicators in serum are direct reflections of metabolism and health in animals. Notably, different breeds of chickens have different growth and metabolism characteristics due to their different genetic backgrounds and thus show different levels of serum biochemical indicators. In the current study, we performed a GWAS of seventeen serum biochemical indicators in a total of 734 F<sub>2</sub> chickens from Gushi chicken × Anka chicken populations to identify markers and candidate genes that affect serum biochemical indicators, laying a foundation to exploiting genomic information for commercial breeding of chickens.

## Results

### Phenotypes and phenotypic correlations

The descriptive statistics for the serum biochemical indicators are listed in Table S1. The variation was large for all serum biochemical indicators in the F<sub>2</sub> resource population. The maximum coefficient of variation was 129.93% (for creatinine (CREA)), and the minimum was 13.17% (for albumin (ALB)). The pairwise correlation coefficients for the serum biochemical indicators (Table S2) were calculated and clustered (Fig. 1), and the results showed that ALB, TP, and globulin (GLO) were clustered



**Fig. 1** Pearson correlations between phenotypes. The Pearson correlation coefficients of pairs of traits were calculated, and the traits were clustered based on the correlation coefficients. The colors (numbers) represent the pairwise correlation coefficients of the serum biochemical indicators. Red indicates a positive correlation, and blue indicates a negative correlation

together and strongly positively correlated, as were LDL, CHO, and HDL. In contrast, traits such as LDH, TP, GLO and TG were negatively correlated with GLU.

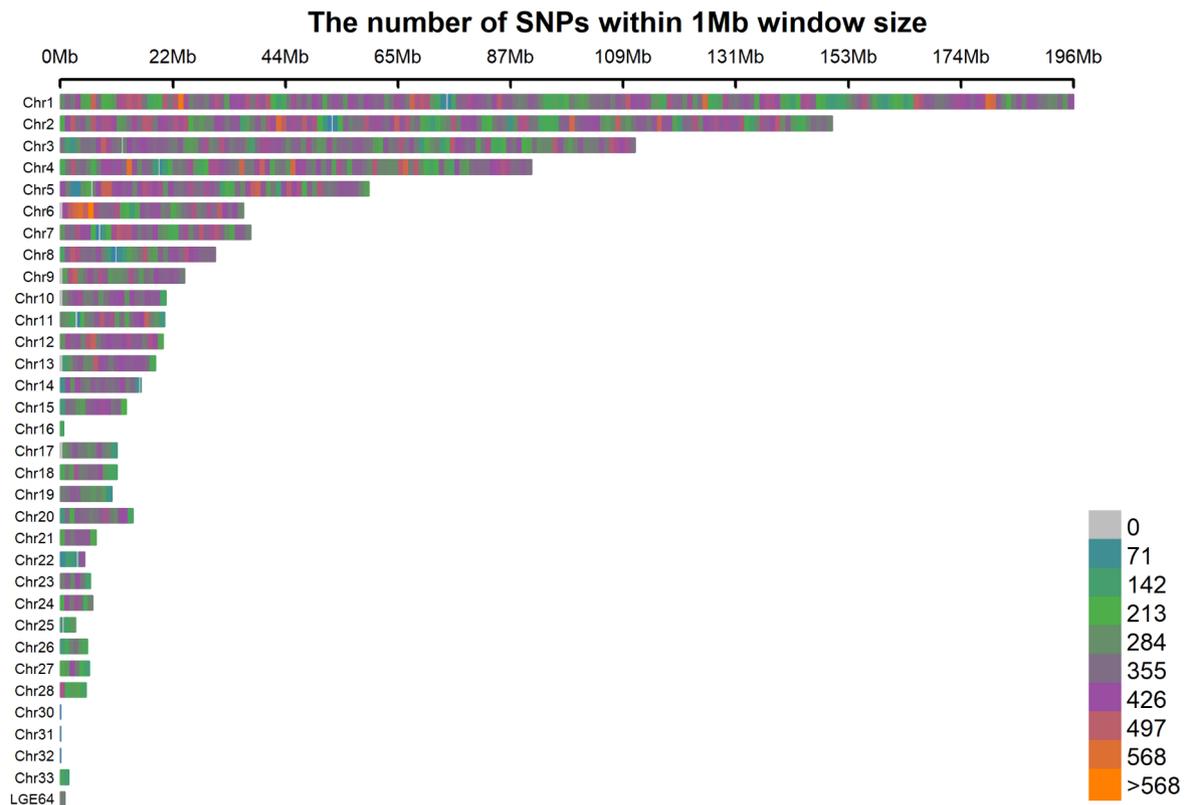
**Summary of genotyping results**

A total of 768 samples were genotyped by sequencing, which yielded 7.258 billion clean sequencing reads and 6.071 billion good barcode reads. The data quality parameters (Q20 and Q30) satisfied the requirements for the subsequent analyses (Table S3)<sup>[16]</sup>. After a series of quality control steps and removal of sex chromosomes, 734 chickens and 321,314 SNPs remained. Among these SNPs, ~10.59% (35,677 SNPs) were first identified by BLAST analysis of the NCBI chicken SNA database (dbSNP) (Table S4)<sup>[16]</sup>. The average SNP density and average SNP variant rate per chromosome were 309 SNPs/Mb and 5.79 kb/SNP, respectively (Fig. 2). These SNPs

were evenly distributed on the genome (Fig. 3), and their distribution was not significantly related to GC content and repeat sequence distribution on the genome. Annotation analysis of the obtained SNPs, majority of the SNPs identified were located in the intergenic region (26.89%) and intron region (41.66%), and the proportion of SNPs in the exon region was 1.09% (Table S5 and Fig. 4).

**GWAS results for the 17 serum biochemical indicators**

A GWAS was carried out on 17 serum biochemical indicators in 734 individuals from the Gushi×Anka F<sub>2</sub> population with an MLM using GCTA. The results of the GWAS are shown in Fig. 5 and Figure S1. The single-marker analysis identified 236 SNPs significantly associated with 8 of the 17 traits with genome-wide significance (-log<sub>10</sub>(P)>5.72), which were mapped to GGA1, GGA3, GGA5, GGA8, GGA9, GGA15, GGA17, GGA21 and



**Fig. 2** Density of SNPs in the chicken genome

GGA24 (Table S6). The genomic control inflation factor ( $\lambda$ ) calculated for all 17 serum biochemical indicators ranged from 0.98 to 1.01, which indicated mild acceptable false positives except for CREA and ALT (Table S7).

**Loci for the AKP trait**

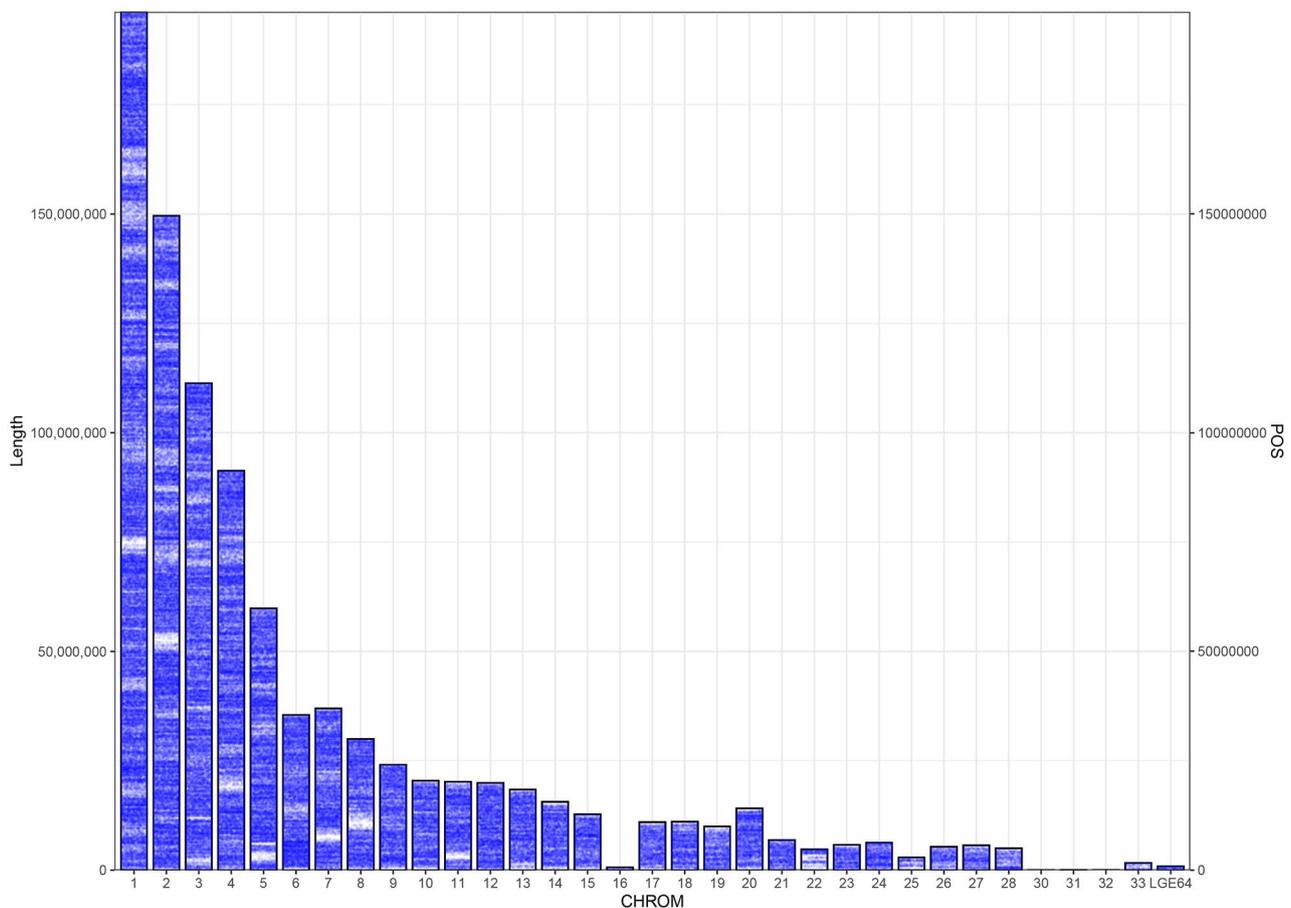
Thirty-one SNPs significantly associated with the AKP trait were identified. A Manhattan plot of the AKP association results is presented in Fig. 5. These 31 SNPs were distributed on GGA3, GGA21 and GGA24, with peaks on GGA21 (5.46–6.67 Mb) and GGA24 (0.27–2.27 Mb). A total of 16 known genes were annotated, including *FAM49A*, *CD30*, *GUCA2A*, *LACTBL1*, *EPHB2*, *CIQA*, *EPHA8*, *PPIH*, *uc\_338*, *ROBO3*, *DCPS*, *KIRREL3*, *BARX2B*, *APLP2*, *ZBTB44* and *OPCML*. The most significant SNP markers on GGA24 affecting AKP were rs14292105 and rs14292106, located downstream of the *DCPS* gene ( $-\log_{10}(P)=12.21$ ). Detailed descriptions of the significant SNPs are given in Table S6.

**Loci for the CHE trait**

Sixty-eight SNPs were identified to be significantly associated with CHE (Fig. 5 and Table S6); these SNPs were located within a 4.7 Mb (19.0–23.7 Mb) region with a peak on GGA9. Twenty known genes were annotated within the region, including *FNDC3B*, *TNIK*, *EIF5A2*, *PRKCI*, *GPR160*, *LRRC34*, *MECOM*, *BCHE*, *OTOL1*, *PPM1L*, *IQCJ*, *MFSD1*, *RSRC1*, *VEPH1*, *CCN1*, *SSR3*, *RAP2B*, *ARHGEF26*, *MINDY4B* and *EIF2A*. The most significant SNP marker affecting CHE was located at the intergenic region of the *ENSGALG00000009511* and *ENSGALG00000030029* genes.

**Loci for the GGT trait**

A total of 130 SNPs spanning a 6.18 Mb region between 6.51 (rs14020787) and 12.69 (rs314303980) Mb on GGA15 were significantly associated with the GGT trait (Fig. 5 and Table S6). Forty-eight known genes were annotated for these SNPs. Interestingly, GGT2 and GGT5 genes were not included in these 48 genes due to



**Fig. 3** Distribution of SNPs in the chicken genome

the lack of significant SNPs at these loci. However, we found that both *GGT2* and *GGT5* genes were within the above mentioned 6.18 Mb region.

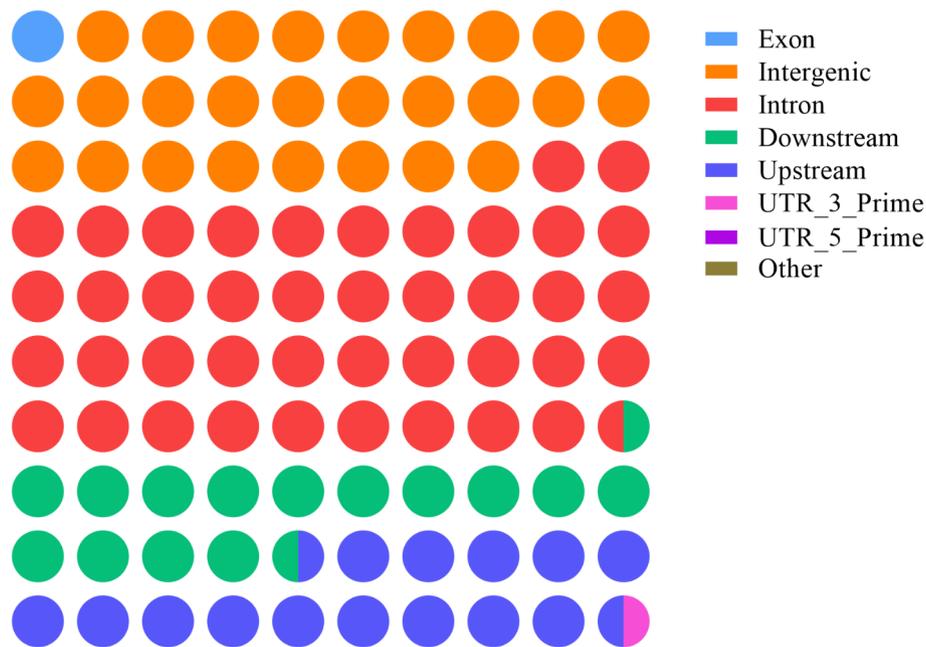
#### Significant loci for the other five traits

GWAS analyses were also performed for other traits, and significant SNPs were identified for five of them, including *AMY*, *LDH*, *GLU*, *HDL* and *TP*. Three SNPs identified at the 14.1 Mb site of *GGA8* were found to be associated with *AMY*; these SNPs were all located within the intergenic region of the *RPL5* and *ENSGALG00000029582* genes. One SNP (rs736356742,  $-\log_{10}(P)=6.19$ ) marker that was significantly associated with *LDH* was located in the intergenic region between the *PTPN5* and *PTPRJ* genes at the 12.8 Mb site of *GGA5*. One SNP marker (rs315358973,  $-\log_{10}(P)=5.91$ ) for *GLU* was identified on *GGA17* downstream of the *TNC* gene. One SNP marker (rs794372306,  $-\log_{10}(P)=6.07$ ) for *HDL* was identified on *GGA1* within the intergenic region of the *AQP11* and *PAK1* genes. One SNP (rs734055122,  $-\log_{10}(P)=7.55$ ) located within the Coatomer Protein Complex Subunit Gamma 2 (*COPG2*) gene was identified to be associated with *TP*.

#### Discussion

Serum biochemical indicators are blood parameters linked to physiological traits and are often regarded as direct reflections of metabolism and health in animals. To the best of our knowledge, no published study has systematically demonstrated the associations between some or all of the 17 serum biochemical indicators and genomic loci using a GWAS of SNPs in chickens. QTLs related to 8 of the 17 indicators (*HDL*, *AKP*, *CHE*, *GGT*, *CREA*, *ALT*, *AST* and *AMY*) in chickens have yet to be reported in the Chicken QTL Database (<https://www.animalgenome.org/cgi-bin/QTLdb/GG/index>). In our study, we identified significant associations for 8 traits, including *AKP*, *CHE*, *GGT*, *AMY*, *LDH*, *GLU*, *HDL* and *TP*. In addition, we performed an analysis of phenotypic correlation to use GWAS find genetic-based associations for some strongly correlated phenotypics. Unfortunately, none of the strongly correlated phenotypic, such as *ALB*, *TP*, and *GLO*, identified significantly associated SNP in the GWAS analysis.

*AKP* is a family of enzymes exists in almost all animal tissues that catalyse the hydrolysis phosphate esters. *AKP* in serum mainly originates in the skeleton and liver

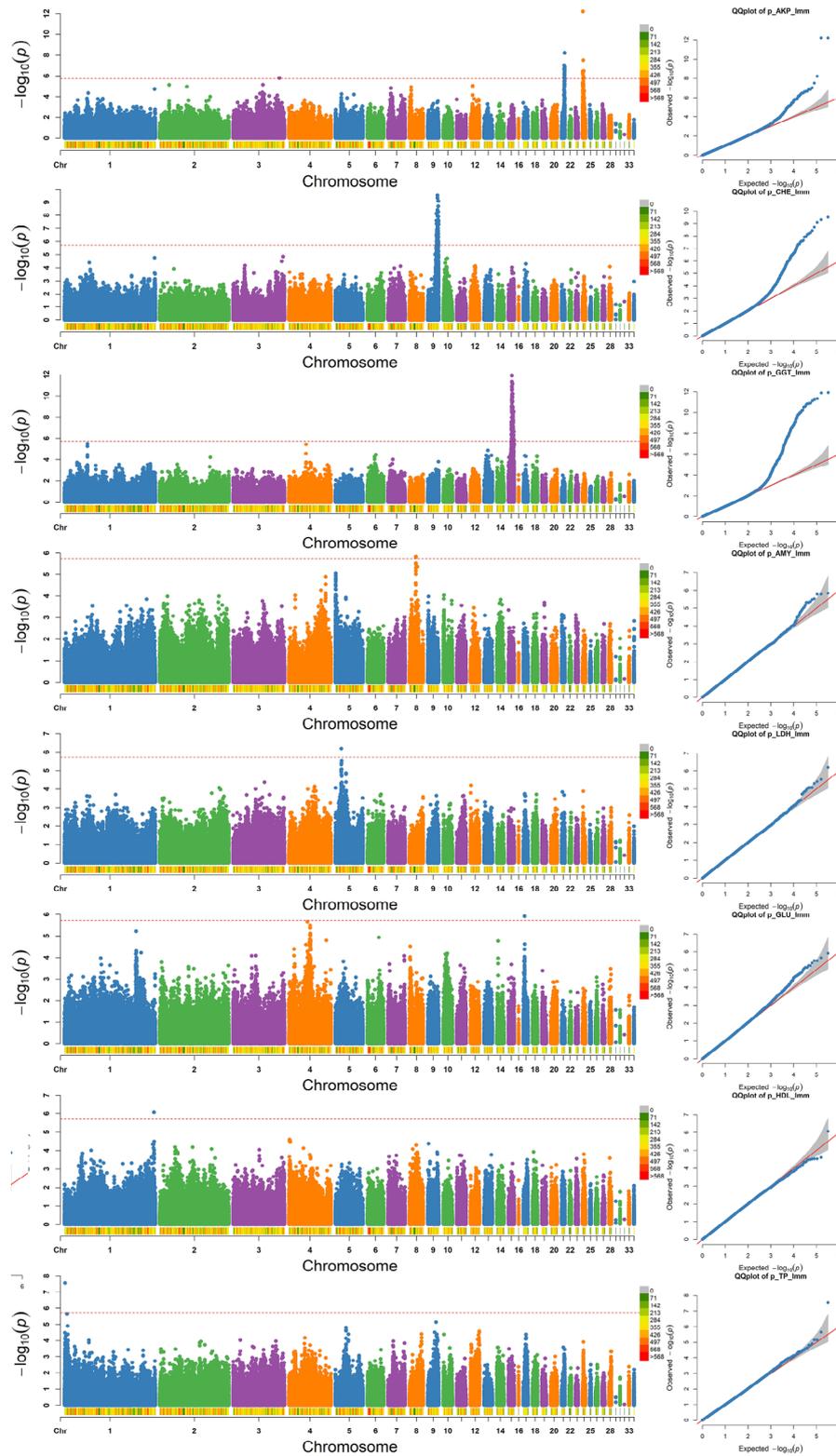


**Fig. 4** Annotated results of SNPs from GBS sequencing. A circle represents 1% of SNPs, and the color of the circle represents SNPs annotated to various functional regions

which often used to diagnose and monitor skeletal or liver disease in the clinic. The levels of AKP in serum is under considerable genetic control according to the Gene variants and heritabilities in swine and human [16, 17]. However, QTL influencing AKP in chicken have not yet been reported. In this study, three QTLs for the AKP trait were identified on GGA3, GGA21 and GGA24, which included sixteen genes suggests that AKP in chicken also under considerable genetic control. Variation within the *ALPL* gene is related to the variation of human serum AKP and bone mineral density [18]. The *ALPL* gene has been proposed as a candidate gene for the AKP trait in pigs and humans [19, 20]. We found that the *ALPL* gene was not directly annotated in our results, but it was located within the QTL on GGA21 which corresponds to a QTL for liver/bone/kidney-derived AKP (ALPL) serum activity. Therefore, a variant or linked variant within the *ALPL* gene may be responsible for the QTL effect on GGA21 in cis-acting substances in this study. In this study, another QTL related to AKP trait were identified on GGA24. The QTL was also identified in the case-control-based GWAS study of the AKP trait mentioned in the background with VVD broilers and sound broilers, the candidate genes located in the QTL were involved in the skeletal development and bone disease. Genetic variation loci of *ST3GAL4* gene located in the QTL were closely related to AKP in VVD broilers in their subsequent study which indicated *ST3GAL4* plays an important role in serum AKP and the skeletal development [3].

Other wise, QTLs associated with AKP activity reported before were mainly involved in growth and bone formation [20, 21]. We postulate that *ST3GAL4* or other genes within the QTL may have a direct or indirect regulatory relationship with the *ALPL* gene to influence the level of AKP in serum and the skeletal development in chickens. Of course, further evidence is needed to understand the regulatory mechanisms. Moreover, with no QTL on GGA9, the location of the intestinal-derived AKP (*ALPI*) gene, the corresponding ALPI enzyme does not seem to be involved in AKP variation in F2 chickens in this study.

CHE is an esterase that can break down choline esters and its activity is related to a family of glycoprotein. There are two types of CHE in vertebrates, Acetylcholinesterase (AChE) and butylcholinesterase (BChE). The former is found in the postsynaptic membrane at the neuromuscular junction, while the latter is synthesized in the liver and has been reported may play a role in lipid metabolism [22]. The activity of CHE in serum is dominated by BChE, the enzyme is often used as an indicator of liver status disease in medical and veterinary trials. Although serum CHE is widely used as a clinical indicator, there are few reports on related QTLs. A previous GWAS on humans has shown that three independent loci on or near the *BCHE* gene are significantly associated with CHE activity in serum [23]. In ducks, Zhu et al [21]. identified 7 SNPs near the *BCHE* gene were significant associations with BChE activity. Consistent with these findings, a QTL for the CHE trait on GGA9



**Fig. 5** Manhattan plots (left) and Q-Q plots (right) for the 8 traits with significantly associated SNPs. Each dot in this figure corresponds to a SNP within the dataset. In each Manhattan plot, the dot color indicates the chromosome on which the SNP is located, the dot position indicates the  $-\log_{10}$ -transformed  $P$  value of the SNP, the number below represents the chromosome number, the length of the figure above the number represents the length of the chromosome, and the color represents the number of SNPs on the chromosome. The horizontal red dashed line denotes the genome-wide significance ( $-\log_{10}(P)=5.72$ ). For each Q-Q plot, the x-axis represents the expected  $-\log_{10}$ -transformed  $P$  value, the y-axis shows the observed  $-\log_{10}$ -transformed  $P$  value, and the red line is the diagonal line

was detected in our study, and the *BCHE* gene within the QTL was also detected to be significantly associated with CHE activity. The above studies of different species all show that CHE activity in serum is significantly associated with *BCHE* gene, which further confirmed that CHE activity in serum is mainly BCHE synthesized in the liver. We further annotated 19 other genes in this QTL on GGA9, while there are few detailed reports on the relationship between CHE and these genes. The most plausible hypothesis is that one or several genes within this QTL might have direct or indirect regulatory effects on *BCHE* gene expression and then affect CHE activity. More molecular studies are needed to reveal the underlying mechanisms of CHE activity.

GGT is a hepatobiliary enzyme synthesized by hepatocytes and intrahepatic bile duct epithelial cells, and its elevated activity is an important predictor of metabolic syndrome. Elevated GGT is a sign of liver disease, and continuously elevated GGT increases the risk of fatty liver disease [24]. Notably, no QTLs for the GGT trait have been reported in any species, thus, our study may be the first to report QTLs that affect GGT level. The QTLs we identified were located within a 6.18 Mb region (6.51~12.69 Mb) on GGA15, and the *GGT2* and *GGT5* genes were located within this QTL. *GGT2* and *GGT5* is members of the  $\gamma$ -glutamyl transpeptidase gene family which including *GGT1*, *GGT6*, *GGT7* and so on. The QTL for the GGT activity in serum was identified in the genomic region where *GGT2* and *GGT5* genes were located, and it was speculated that the GGT activity in serum was mainly regulated by these two genes.

One significant SNP located at the 12.8 Mb site of GGA5 was identified to be associated with the LDH trait, and the L-lactate dehydrogenase A chain (*LDHA*) that belongs to the lactate dehydrogenase family was annotated to a site 0.2 Mb upstream of the 12.8 Mb site on GGA5. Gene Ontology annotations related to *LDHA* include oxidoreductase activity and L-lactate dehydrogenase activity. Therefore, a variant or linked variant within the *LDHA* gene may be responsible for the QTL effect on GGA5 associated with the LDH trait in this study and the *LDHA* gene could be considered the site of a candidate gene.

The GLU trait was mapped to GGA17 in our study. A previous study has mapped GLU traits elsewhere on GGA17[25]. In addition, a total of 23 QTLs on different chromosomes have been associated with GLU levels in studies based on different experimental designs [13, 25–30]. For example, Hee-Bok Park et al. performed QTL analysis on an F2 population with 874 individuals composing 75 full-sib families and found that one QTL on GGA20 at the 41 Mb site and one QTL on GGA27 at the 21 Mb site were related to the GLU trait [28]. The *TNC* gene was also annotated to a locus we detected; this

gene has been demonstrated to be highly associated with a wide range of diseases related to inflammation, including diabetes [31]. In addition, the *TNC* gene is involved in the regulation of adipocyte differentiation and lipid deposition, and the expression of the porcine *TNC* gene is associated with meat quality [32]. Fat is an important source of stored energy, while sugar mainly provides energy for metabolism; however, fat and sugar can be interconverted. The *TNC* gene may affect GLU levels by regulating the relationship between fat deposition and sugar metabolism and thus may be a candidate gene for the GLU trait.

HDL is an antiatherosclerotic plasma lipoprotein that removes cholesterol from atherosclerotic blood vessels and transports it to the liver for metabolic clearance [33]. In our study, the HDL trait was mapped to the region on GGA1 containing the *AQP11* and *PAK1* genes. To our knowledge, no QTLs on GGA1 that are associated with HDL levels in chickens have been reported thus far; however, the interval in which the *AQP11* and *PAK1* genes are located in the pig genome is also associated with HDL levels [34]. *PAK1* plays important roles in a variety of cellular processes, including cell polarity, motility, survival and proliferation [35]. The increase in PAK1 phosphorylation in human atherosclerotic arteries suggests a role for this gene in human atherogenesis [35]. The *PAK1* gene also promotes activation of inflammatory pathways in endothelial cells by fluid shear stress, which is the initial stimulator of atherosclerosis [36]. Transport of cholesterol via HDL from peripheral tissues to the liver is a mechanism associated with very advanced antiatherosclerosis [33]. *PAK1* mediates the downregulation of scavenger receptor B1 promoter activity to play a role in cholesterol transport [37]. Notably, *PAK1* deficiency leads to enhancement of cholesterol efflux and upregulation of reverse cholesterol transporters, which suggests that *PAK1* exerts negative modulatory effects on transporters and may promote lipid retention in inflamed arteries, causing atherogenesis [35]. Therefore, *PAK1* may be a candidate gene affecting serum levels of HDL.

The serum AMY trait was mapped to the 14.1 Mb site of GGA8, 3.1 Mb downstream of the *AMY1A* gene. The *AMY1A* gene encodes AMY, a secreted protein that hydrolyzes 1,4- $\alpha$ -glucoside bonds in oligosaccharides and polysaccharides and thus catalyzes the first step in the digestion of dietary starch and glycogen. Genetic variation in this genomic interval may affect the expression of *AMY* genes or the translation of proteins through indirect effects, thereby regulating the levels of AMY in serum.

TP is an indicator of growth and nutritional status [38]. The level of serum TP of fat pigs are higher than that of lean pigs [9], and the level of serum TP has been proved to be the best indicator for early fat estimation in

pigs [10]. In addition, TP can also be used as a potential biomarker to select the fat traits of ducks [21]. Previous research has identified six QTLs related to the TP trait that are located on GGA1, GGA2, GGA12, GGA18 and GGAZ [25]; however, our study has identified a new QTL on GGA1 that is located within the *COPG2* gene. The coatomer is a cytosolic protein complex that further mediates biosynthetic protein transport from the endoplasmic reticulum (ER) via the Golgi to the trans-Golgi network. The *COPG2* gene was enriched for the Delta508-CFTR traffic/ER-to-Golgi in cystic fibrosis (CF) pathway, vesicle-mediated transport pathway, and protein metabolism pathway, among other pathways (<https://www.genecards.org/cgi-bin/carddisp.pl?gene=COPG2&keywords=COPG2>). Thus, *COPG2* may regulate the TP trait by affecting protein transport and metabolism.

In summary, although the QTLs and candidate genes identified by our GWAS have not been reported in other studies, we have confirmed the associations of the QTLs/candidate genes with serum biochemical indicator traits in other species or the biological functions of the potential candidate genes. The different association results among different studies might have been caused by the different experimental populations, different genetic backgrounds and different stages at which the traits were measured.

## Conclusion

This study identified 10 novel QTL regions that affected 8 serum biochemical indicators based on a Gushi×Anka F<sub>2</sub> resource population using genotyping by sequencing. We propose that the *ALPL*, *BCHE*, and *GGT2/GGT5* genes might affect the AKP, CHE and GGT traits, respectively. The results provide a basis for uncovering the molecular mechanism of serum biochemical indicator regulation in chicken.

## Methods

### Establishment of the experimental population

The F<sub>2</sub> resource population used in this study was created by the Henan Innovative Engineering Research Center of the Poultry Germplasm Resource as described previously [39]. The F<sub>2</sub> resource population was established with a slow-growing breed (the Gushi chicken) and a fast-growing breed (the Anka chicken), which consisted of 4 cross-bred families (Anka cocks mated with Gushi hens) and 3 reciprocal families (Gushi cocks mated with Anka hens) as described previously [39]. A total of 860 chickens were obtained in the F<sub>2</sub> population. All chickens were kept in cages under the same environmental conditions with free access to feed and water according to the National Research Council recommendations (1994).

### Phenotypic measurements

Blood samples were collected for all 860 F<sub>2</sub> chickens after the chickens were slaughtered at 12 weeks of age. Then, serum samples were separated by a standard protocol and stored at -80 °C for measurement of serum biochemical indicators. A total of 17 serum biochemical indicators, including CHO, TGs, HDL, LDL, glucose (GLU), AKP, CHE, creatine phosphokinase (CK),  $\gamma$ -glutamyl transpeptidase (GGT), lactate dehydrogenase (LDH), CREA, TP, GLO, ALB, ALT, AST, and amylase (AMY), were measured with kits from Nanjing Jiancheng (Nanjing, China). The measurement methods have been previously detailed by Han et al. [40, 41]. All phenotypic data were analyzed with normality tests. If a phenotype did not follow the normal distribution (asymptotic  $P < 0.05$ ), it would be ranked as normal scores with the Tukey method, and then transformed data were used in the following genetic analyses. The pheatmap R package (1.0.12) was applied to visualize the Pearson correlations between phenotypes.

### Genotyping and quality control

Genomic DNA was extracted from blood samples with a DNeasy Blood and Tissue Kit according to the manufacturer's instructions (Qiagen, Hilden, Germany). All DNA samples were diluted to 40 ng/ $\mu$ l, and 200 ng of DNA was digested with a combination of EcoRI and MseI for double-digest genotyping by sequencing (ddGBS) [42]. Eight libraries were constructed, and a total of 768 samples were genotyped. Sequencing experiments were performed via paired-end 2×150 nt runs on an Illumina HiSeq X Ten platform controlled by data collection software. The image data were exported, transformed into raw data and stored in FASTQ (fq) format after sequencing. The 150 bp paired-end Illumina reads containing the adapter sequence were deleted, and those containing more than 50% low-quality bases or more than 5% N bases were removed. A quality control report for the filtered reads was generated by FastQC software. All downstream analyses were based on the resulting clean data.

The TASSEL-GBS analysis pipeline (version 5.2.31) [43, 44] was used to call single-nucleotide polymorphisms (SNPs), and the reads were aligned to the chicken reference genome *Gallus\_gallus*-5.0 (*Gallus gallus* 5.0) using Bowtie2 (version 2.3.0) [45]. Then, we used VCFtools (version 0.1.13) [46] to filter the raw SNPs with the following parameters: a minor allele frequency (MAF) greater than 5% (maf 0.05), a genotype quality greater than 98 (minGQ 98), a genotype depth greater than 5 (minDP 5), retention of only biallelic markers (max-alleles 2; min-alleles 2), conformation to Hardy-Weinberg equilibrium (hwe 0.0001) and a maximum missing rate less than 0.4. Individual samples were excluded with two-sided call rates of less than 0.3. We merged the SNPs of paired-end reads and imputed the ungenotyped markers

according to the information on the remaining SNPs with Beagle4.0 software [47]. SNPs located on sex chromosomes (GGAZ and GGAW) were removed before the GWAS was performed. We have analyzed the population structure and removed the abnormal individuals in previous study [48].

### Single-trait GWAS

The top two principal components (PCs) were used as covariates in the mixed model [48]. A genomic relationship matrix was constructed with the SNPs using GCTA software and used as a random effect in the mixed model.

A GWAS was carried out on serum biochemical indicators using the following mixed linear model (MLM) equation in the GCTA program:

$$y = X\beta + X_{SNP}\beta_{SNP} + Zu + e$$

where  $y$  is the vector of individual observations,  $\beta$  is the vector of fixed environmental effects (sex, batch and top two PCs),  $u$  is the vector of random direct additive genetic effects obtained from MVN ( $u \sim 0, \sigma^2 u$ ), and  $e$  is the vector of random residual errors obtained from MVN ( $e \sim 0, \sigma^2 e$ ). The  $X$  and  $Z$  are design matrices that relate individuals' records to their fixed effects ( $\beta$ ) and additive genetic effects ( $u$ ), respectively.  $X_{SNP}$  is the incidence matrix for the SNP markers and  $\beta_{SNP}$  is the regression coefficient for each SNP (SNP effects).

The genome-wide significance ( $P$  value threshold) was adjusted with the Bonferroni method based on the number of independent tests. We calculated the number of genome-wide independent markers using the PLINK [49] `--indep-pairwise` command with a window size of 25 SNPs, a step of 5 SNPs, and an  $r^2$  threshold of 0.1. The significance level was set as  $1.90E-06$  ( $0.05/26,352$ ;  $-\log_{10}(P) > 5.72$ ) based on the 26,352 independent markers. Manhattan and quantile-quantile (Q-Q) plots were obtained using the CMplot package (<https://github.com/YinLiLin/R-CMplot>) within R software (version 3.6.1) (<http://www.r-project.org/>).

### Post-GWAS analysis

The Ensembl genome database and the SNPEff (version 4.1) program [50] were used to obtain information about SNPs and relevant gene annotations for the GWAS results. Additionally, the plotrix (3.7-8), lattice-Extra (0.6–29), scales (1.1.0) and ggplot2 (3.3.0) packages within R software (version 4.0.0) were applied to visualize the distributions of chromosome length and SNPs in the chicken genome. The CMplot package was used to visualize the density of SNPs in the chicken genome.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-023-09206-7>.

**Additional file 1. Table S1.** Descriptive statistics for the serum biochemical indicators

**Additional file 2. Table S2.** Pearson correlation coefficients between serum biochemical indicators

**Additional file 3. Table S3.** Statistics for the sequenced data

**Additional file 4. Table S4.** Distribution of SNPs discovered from 734 individuals across chromosomes

**Additional file 5. Table S5.** Annotated results of SNPs from GBS sequencing

**Additional file 6. Table S6.** Significant SNPs for serum biochemical indicators.

**Additional file 7. Table S7.** The genomic inflation factor ( $\lambda$  statistic) for serum biochemical indicators

**Additional file 8. Figure S1**

### Acknowledgements

Not applicable.

### Author Contribution

XK, YT and YZ: conceived and designed the experiments. HS, YZ and WL performed the data analysis, wrote the manuscript and prepared the figures. XK, YT, YL, BZ, RH, GS, RJ, ZL, FY, and GL constructed the F2 population and collected the phenotypes. YG, YC, and XL prepared the samples, figures and tables. All authors read and approved the final manuscript.

### Funding

This work was supported by the Scientific Studio of Zhongyuan Scholars (30601985), Zhongyuan Science and Technology Innovation Leading Scientist Project (214200510003), the National Natural Science Foundation of China (31902144), Program for Innovative Research Team (in Science and Technology) in University of Henan Province (21IRTSTHN022), and China Agriculture Research System of MOF and MARA (CARS-40-K04).

### Data Availability

The authors declare that the data supporting the findings of this study are available within the article and its supplementary information files. All the raw sequences have been deposited in the NCBI database Sequence Read Archive with the accession numbers SRR12532401–SRR12532408 (BioProject number PRJNA659316).

### Declarations

#### Ethics statement

All animal experiments were performed according to the Regulations for the Administration of Affairs Concerning Experimental Animals (Ministry of Science and Technology, China, 2004). The protocols and guidelines were approved by the Institutional Animal Care and Use Committee of Henan Agricultural University, China. The study was carried out in compliance with the ARRIVE guidelines.

#### Consent for publication

Not applicable.

#### Competing Interest

The author declares that no conflict of interest could be perceived as prejudicing the impartiality of the research reported.

Received: 2 July 2022 / Accepted: 23 February 2023

Published online: 02 March 2023

## References

1. Millán JL. Alkaline phosphatases: structure, substrate specificity and functional relatedness to other members of a large superfamily of enzymes. *Purinergic Signalling*. 2009;2(2):335.
2. Wu SJ, Lin YX, Ye H, Xiong XZ, Li FY, Cheng NS. Prognostic value of alkaline phosphatase, gamma-glutamyl transpeptidase and lactate dehydrogenase in hepatocellular carcinoma patients treated with liver resection. *Int J Surg*. 2016;36(Pt A):143–51.
3. Guo Y, Huang H, Zhang Z, Ma Y, Li J, Tang H, Ma H, Li Z, Li W, Liu X, et al. Genome-wide association study identifies SNPs for growth performance and serum indicators in Valgus-varus deformity broilers (*Gallus gallus*) using ddGBS sequencing. *BMC Genomics*. 2022;23(1):26.
4. Ogunkeye OO, Roluga AI. Serum cholinesterase activity helps to distinguish between liver disease and non-liver disease aberration in liver function tests. *Pathophysiology*. 2006;13(2):91–3.
5. Ray W, Flamm K, Bisceglie SL, Di AM, Bodenheimer HC. Serum activity of alanine aminotransferase (ALT) as an indicator of health and disease. *Hepatology*. 2008;47(4):1363–70.
6. Kayamori F, Igarashi K. Effects of Dietary Nasunin on the serum cholesterol level in rats. *J Agricultural Chem Soc Japan*. 1994;58(3):570–1.
7. Benson MK, Devi K. Influence of omega-6/omega-3 rich dietary oils on lipid profile and antioxidant enzymes in normal and stressed rats. *Indian J Exp Biol*. 2009;47(2):98–103.
8. Guo L, Sun B, Shang Z, Leng L, Wang Y, Wang N, Li H. Comparison of adipose tissue cellularity in chicken lines divergently selected for fatness. *Poult Sci*. 2011;90(9):2024–34.
9. He Q, Ren P, Kong X, Wu Y, Wu G, Li P, Hao F, Tang H, Blachier F, Yin Y. Comparison of serum metabolite compositions between obese and lean growing pigs using an NMR-based metabolomic approach. *J Nutr Biochem*. 2012;23(2):133–9.
10. Muñoz R, Tor M, Estany J. Relationship between blood lipid indicators and fat content and composition in Duroc pigs. *Livest Sci*. 2012;148(1–2):95–102.
11. Kao WH, Folsom AR, Nieto FJ, Mo JP, Watson RL, Brancati FL. Serum and dietary magnesium and the risk for type 2 diabetes mellitus: the atherosclerosis risk in Communities Study. *Arch Intern Med*. 1999;159(18):2151–9.
12. Wittman JC, Grobbee DE, Derckx FH, Bouillon R, de Bruijn AM, Hofman A. Reduction of blood pressure with oral magnesium supplementation in women with mild to moderate hypertension. *Am J Clin Nutr*. 1994;60(1):129–35.
13. Nadaf J, Pitel F, Gilbert H, Duclos M, Vignoles F, Beaumont C, Vignal A, Porter T, Cogburn L, Aggrey S. QTL for several metabolic traits map to loci controlling growth and body composition in an F<sub>2</sub> intercross between high- and low-growth chicken lines. *Physiol Genom*. 2009;38(3):241–9.
14. Park HB, Jacobsson L, Wahlberg P, Siegel PB, Andersson L. QTL analysis of body composition and metabolic traits in an intercross between chicken lines divergently selected for growth. *Physiol Genomics*. 2006;25(2):216–23.
15. Zhou H, Evock-Clover C, McMurtry J, Ashwell C, Lamont SJ. Genome-wide linkage analysis to identify chromosomal regions affecting phenotypic traits in the chicken. IV. Metabolic traits. *Poult Sci*. 2007;86(2):267–76.
16. Reiner G, Hepp S, Hertrampf B. Genetisch determinierte Varianzanteile klinisch-chemischer Laborparameter beim Schwein. *Tierärztliche Praxis Ausgabe G: Großtiere / Nutztiere*. 2006;34:40–9.
17. Havill LM, Rogers J, Cox LA, Mahaney MC. QTL with pleiotropic effects on serum levels of bone-specific alkaline phosphatase and osteocalcin maps to the baboon ortholog of human chromosome 6p23-21.3. *J Bone Miner Res*. 2006;21(12):1888–96.
18. Goseki-Sone M, Sogabe N, Fukushi-Irie M, Mizoi L, Orimo H, Suzuki T, Nakamura H, Orimo H, Hosoi T. Functional analysis of the single nucleotide polymorphism (787T > C) in the tissue - nonspecific alkaline phosphatase gene Associated with BMD. *J Bone Miner Res*. 2005;20(5):773–82.
19. Yoo CK, Cho IC, Lee JB, Jung EJ, Lim HT, Han SH, Lee SS, Ko MS, Kang T, Hwang JH. QTL analysis of clinical-chemical traits in an F<sub>2</sub> intercross between Landrace and Korean native pigs. *Physiol Genom*. 2012;44(13):657–68.
20. Reiner G, Clemens N, Fischer R, Köhler F, Berge T, Hepp S, Willems H. Mapping of quantitative trait loci for clinical-chemical traits in swine. *Anim Genet*. 2009;40(1):57–64.
21. Zhu F, Cui QQ, Yang YZ, Hao JP, Yang FX, Hou ZC. Genome-wide association study of the level of blood components in Pekin ducks. *Genomics*. 2020;112(1):379–87.
22. Valle A, O'Connor DT, Taylor P, Zhu G, Montgomery GW, Slagboom PE, Martin NG, Whitfield JB. Butyrylcholinesterase: association with the metabolic syndrome and identification of 2 gene loci affecting activity. *Clin Chem*. 2006;52(6):1014–20.
23. Beben B, Middelberg RP, Lind PA, Valle AM, Scott G, Nyholt DR, Medland SE, Henders AK, Heath AC, Madden PAF. GWAS of butyrylcholinesterase activity identifies four novel loci, independent effects within BCHE and secondary associations with metabolic risk factors. *Hum Mol Genet*. 2011;20(22):4504–14.
24. Fujii H, Doi H, Ko T, Fukuma T, Kadono T, Asaeda K, Kobayashi R, Nakano T, Doi T, Nakatsugawa Y, et al. Frequently abnormal serum gamma-glutamyl transferase activity is associated with future development of fatty liver: a retrospective cohort study. *BMC Gastroenterol*. 2020;20(1):217.
25. Javanrouh-Aliabad A, Torshizi RV, Masoudi AA, Ehsani A. Identification of candidate genes for blood metabolites in Iranian chickens using a genome-wide association study. *Br Poult Sci*. 2018;59(4):381–8.
26. Demeure O, Jean S, Bihan-Duval L, Elisabeth. Pitel, Frederique, Boland, Anne: genome-wide interval mapping using SNPs identifies new QTL for growth, body composition and several physiological variables in an F<sub>2</sub> intercross between fat and lean chicken lines. *Genet Selection Evol Gse*. 2013;45(1):36.
27. Huang Y, Zhang C, Zhang W, Zhang P, Kang X, Chen W. Variation in the chicken LPIN2 gene and association with performance traits. *Br Poult Sci*. 2015;56(2):175–83.
28. Hee-Bok P, Lina J, Per W, Siegel PB, Leif A. QTL analysis of body composition and metabolic traits in an intercross between chicken lines divergently selected for growth. *Physiol Genomics*. 2006;25(2):216–23.
29. Goor AV, Ashwell CM, Persia ME, Rothschild MF, Schmidt CJ, Lamont SJ. Quantitative trait loci identified for blood chemistry components of an advanced intercross line of chickens under heat stress. *BMC Genomics*. 2016;17(1):287.
30. Zhou H, Deeb N, Evock-Clover CM, Ashwell CM, Lamont SJ. Genome-wide linkage analysis to identify chromosomal regions affecting phenotypic traits in the chicken. I. Growth and average daily gain. *Poult Sci*. 2007;85(10):1712–21.
31. Victoria C, Javier GA, Amaia R, Beatriz R, Fernando R, Victor V, Camilo S, Gil MJ, Javier S, Gema F. Increased tenascin C and toll-like receptor 4 levels in visceral adipose tissue as a link between inflammation and extracellular matrix remodeling in obesity. *J Clin Endocrinol Metabolism*. 2012;97(10):E1880.
32. Kayan A, Cinar MU, Uddin MJ, Phatsara C, Wimmers K, Ponsuksili S, Tesfaye D, Looft C, Juengst H, Tholen E. Polymorphism and expression of the porcine tenascin C gene associated with meat and carcass quality. *Meat Sci*. 2011;89(1):76–83.
33. Glomset JA. The plasma lecithin: cholesterol acyltransferase reaction. *J Lipid Res*. 1968;9:155–67.
34. Chen R, Ren J, Li W, Huang X, Yan X, Yang B, Zhao Y, Guo Y, Mao H, Huang L. A genome-wide scan for quantitative trait loci affecting serum glucose and lipids in a White Duroc × Erhualian intercross F<sub>2</sub> population. *Mamm Genome*. 2009;20(6):386–92.
35. Singh NK, Kotla S, Dyukova E Jr, Orr TJ, Chernoff AW, Marion J, Rao TN. Disruption of p21-activated kinase 1 gene diminishes atherosclerosis in apolipoprotein E-deficient mice. *Nat Commun*. 2015;6:7450.
36. Jhaveri KA. The role of p21-activated kinase in the initiation of atherosclerosis. *BMC Cardiovasc Disord*. 2012;12(1):55–5.
37. Hullinger TG, Panek RL, Xu X, Karathanasis SK. p21-activated kinase-1 (PAK1) inhibition of the human scavenger receptor class B, type I promoter in macrophages is independent of PAK1 kinase activity, but requires the GTPase-binding domain. *J Biol Chem*. 2001;276(50):46807–14.
38. Filipović N, Stojević Z, Milinković-Tur S, Beer Ljubić B, Zdelar-Tuk M. Changes in concentration and fractions of blood serum proteins of chickens during fattening. *Veterinarski arhiv*. 2007;77(4):319–26.
39. Han RL, Kang XT, Sun GR, Li GX, Bai YC, Tian YD, Huang YQ. Novel SNPs in the PRDM16 gene and their associations with performance traits in chickens. *Mol Biol Rep*. 2012;39(3):3153–60.
40. Han RL, Li ZJ, Li MJ, Li JQ, Lan XY, Sun GR, Kang XT, Chen H. Novel 9-bp indel in visfatin gene and its associations with chicken growth. *Br Poult Sci*. 2011;52(1):52–7.
41. Li W, Liu D, Tang S, Li D, Han R, Tian Y, Li H, Li G, Li W, Liu X, et al. A multiallelic indel in the promoter region of the cyclin-dependent kinase inhibitor 3 gene is significantly associated with body weight and carcass traits in chickens. *Poult Sci*. 2019;98(2):556–65.
42. Wang Y, Cao X, Zhao Y, Fei J, Hu X, Li N. Optimized double-digest genotyping by sequencing (ddGBS) method with high-density SNP markers and high genotyping accuracy for chickens. *PLoS ONE*. 2017;12(6):e0179073.

43. Van Tassel CP, Smith TP, Matukumalli LK, Taylor JF, Schnabel RD, Lawley CT, Haudenschild CD, Moore SS, Warren WC, Sonstegard TS. SNP discovery and allele frequency estimation by deep sequencing of reduced representation libraries. *Nat Methods*. 2008;5(3):247–52.
44. Glaubitz JC, Casstevens TM, Lu F, Harriman J, Elshire RJ, Sun Q, Buckler ES. TASSEL-GBS: a high capacity genotyping by sequencing analysis pipeline. *PLoS ONE*. 2014;9(2):e90346.
45. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9(4):357–9.
46. Petr D, Adam A, Goncalo A, Cornelis AA, Eric B, Mark AD, Robert EH, Gerton L, Gabor TM, Stephen T. The variant call format and VCFtools. *Bioinformatics*. 2011;27(15):2156–8.
47. Browning BL, Browning SR. Improving the accuracy and efficiency of identity-by-descent detection in population data. *Genetics*. 2013;194(2):459–71.
48. Zhang Y, Wang Y, Li Y, Wu J, Wang X, Bian C, Tian Y, Sun G, Han R, Liu X, et al. Genome-wide association study reveals the genetic determinism of growth traits in a Gushi-Anka F(2) chicken population. *Heredity (Edinb)*. 2021;126(2):293–307.
49. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81(3):559–75.
50. Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)*. 2012;6(2):80–92.

### **Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.