Identification and Genetic Analysis of Major Gene ST3GAL4 Related to Serum Alkaline Phosphatase in Chickens

Hehe Tang
    Henan Agricultural University

Yanchao Ma
    Henan Agricultural University

Jianzeng Li
    Henan Agricultural University

Zhenzhen Zhang
    Henan Agricultural University

Wenting Li
    Henan Agricultural University

Hetian Huang
    Henan Agricultural University

Chunxia Cai
    Henan Agricultural University

Lujie Zhang
    Henan Agricultural University

Zhuanjian Li
    Henan Agricultural University

Ruirui Jiang
    Henan Agricultural University

Yadong Tian
    Henan Agricultural University

Guoxi Li
    Henan Agricultural University

Fengbin Yan
    Henan Agricultural University

Donghua Li
    Henan Agricultural University

Yanhua Zhang
    Henan Agricultural University

Guirong Sun
Research

Keywords: Alkaline phosphatase, ST3GAL4, Linkage disequilibrium, Genotypic and allelic frequencies, Genetic parameters, Tissue expression level

Posted Date: September 15th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-850420/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Background: An increasing number of studies have explored disease and growth traits through quantitative trait locus (QTLs) in chickens. Nevertheless, the pathogenic genes underlying the QTL effects remain poorly understood. Alkaline phosphatase (ALP) is a marker of osteoblast maturation and an important indicator of bone metabolism. The change of ALP can reflect the bone metabolism and growth traits of animals to a certain extent.

Results: In this study, we identified a SNP site at ST3GAL4 that found by genome-wide association studies (GWAS) in previous studies, and found another 8 SNPs by DNA sequencing. Interestingly, there are 4 SNPs rs475471G>A, rs475533C>T, rs475621A>G, rs475647C>A were completely linked by haplotype analysis, and selected a tag SNP rs475471G>A to analyze the correlation between this SNP and ALP level in Hubbard leg disease population and an F2 chicken resource population produced by Anka and Gushi chickens, and carried out population genetic analysis in 18 chicken breeds. Association analysis showed that this QTL within ST3GAL4 were highly correlated with ALP level, the mutant individuals with genotype AA had the highest ALP level, followed by GA and GG genotypes. The mutant individuals with genotype AA and GA genotype had higher values for body weight (BW), chest width (CW), body slanting length (BSL), pelvis width (PW) at 4-week, the semi-evisceration weight (SEW), evisceration weight (EW) and Leg weight (LW) of AA and GA genotype also higher. The amplification and typing of 4,852 DNA samples from 18 different breeds and the result shown GG genotype mainly existed in egg-type chickens and dual-type chickens, while AA genotype mainly distributed in commercial broilers and F2 resource population. The individuals of AA genotype had the highest ALP and showed better growth performance. This is the first time to report the causal variant of ST3GAL4 gene that located at chromosome 24 related to chicken serum ALP level. Besides, tissue expression analysis used Cobb broiler showed that there were significant differences between different genotypes in spleen and duodenum.

Conclusions: This study the first time to determine 9 SNPs within ST3GAL4 were related to ALP in chickens, 4 of them were complete linkage sequence variants, which provide useful information on the mutation of ST3GAL4 and could predict the serum ALP level of chicken early and effectively as a potential molecular breeding marker.

Background

Alkaline phosphatase (ALP) is a group of specific hydrolysis isozymes that catalyze the hydrolysis of organic phosphate esters in alkaline environment [1]. ALP has many main physiological functions in the body, which is a marker of osteoblasts maturation and an important indicator of bone metabolism, and is essential in bone mineralization and functions in vascular calcification [2, 3]. Its activity is affected by many factors, ALP is widely distributed in the body, serum ALP derives mostly from bone, liver, kidney and intestine, the content is highest in intestine mucosa and placenta [4, 5]. When bone metabolism is flourishing, osteoblasts are active and ALP secretion increases, which exists around and on the surface of osteoblasts and is easy to release into the blood. Therefore, when bone growth is hyperactive, osteoblasts
proliferate and serum ALP activity increases, which leads to the change of blood ALP concentration, ALP is by-production during bone formation that can reflect bone development state [6, 7]. Studies in humans have shown that ALP is associated with kidney injury and a variety of bone diseases [8, 9], and the additive genetic heritability of serum ALP in mice is approximately 56% [10].

Golgi-resident sialyltransferases of the ST3GAL family are type II membrane enzymes that specifically catalyzes the transfer of α2,3 linkage to termini of N- and O-glycan chains [11, 12]. Six genes (ST3GAL1-ST3GAL6) encoding these sialyltransferases have been identified. Sialyltransferase plays a key role in the glycoprotein and glycolipid glycosylation. It has been found that the abnormal expression of sialyltransferase is closely related to the occurrence and development of tumor and immune system diseases [13].

ST3GAL4 is a member of the sialic acyltransferase family. It has been reported that ST3GAL4 plays an important role in the pathogenesis of osteoarthritis (OA), and is highly expressed in the cartilage of OA patients [14], which can mediate the degradation, apoptosis and proliferation of extracellular mechanism (ECM). In addition, miR-193b was found to have a target relationship with ST3GAL4, and overexpression of ST3GAL4 could affect the salivation of CD44, reduce its ability to bind to lubricant, and mediate the activity of NF-κB pathway [15–18]. In human, SNP at rs11220462 of ST3GAL4 gene was associated with cholesterol (T-CHO) and low density lipoprotein concentrations (LDL-C) in European populations [19]. However, there are few studies on ST3GAL4 gene polymorphism and molecular mechanisms in chickens.

In our previous studies, the abnormal changes of serum calcium (Ca), phosphorus (P), alkaline phosphatase (ALP) and blood lipids level in Valgus-varus deformity (VVD) broilers, and ALP level of VVD broiler was significantly lower [20]. Further, GWAS also found that rs475471G>A at ST3GAL4 was significantly correlated with ALP level in Hubbard VVD leg disease population (unpublished data). In this research, the relationship between ST3GAL4 genetic variation loci and serum ALP was studied. The heritability of ALP was moderate [21], therefore, understanding the relationship between genetic variation and ALP levels may provide new insights for many related diseases caused by abnormal serum ALP.

Results

Screening and identification of mutation sites of ST3GAL4 gene in chicken

The SNP site rs475471G>A were identified at ST3GAL4 gene based on previous GWAS results. (Unpublished data). To verify this SNP, and 9 mutation sites were found. The DNA samples of 468 Hubbard broilers were sequenced by PCR-SBT (Table S1) and sequencing map is shown in Fig 1, and the position of the mutation site in the gene structure is shown in Fig 2.

Linkage disequilibrium analysis of chicken ST3GAL4 gene mutation sites

Haploview software was used for linkage disequilibrium analysis. The D' value of 9 SNPs (rs475435C>T, rs475460C>T, rs475471G>A, rs475533C>T, rs475621A>G, rs475647C>A, rs475695C>T, rs475743A>G and
rs475744A>G) in ST3GAL4 gene were calculated. The results showed that the D' value of pairing of the above SNPs were greater than 0.97 except for the constraint condition that the D' values of rs475435C>T and rs475743A>G were less than 0.8; The calculation of r² value showed that except for rs475435C>T and rs475743A>G, the r² value of pairing between rs475460C>T, rs475471G>A, rs475533C>T, rs475621A>G, rs475647C>A, rs475695C>T and rs475744A>G was greater than 0.9. When the restriction conditions of r² value greater than 0.33 and D' value greater than 0.8 are satisfied between two polymorphic loci, they are considered to be in a strong linkage disequilibrium state. Through the analysis of the results, we found that the D' value and r² value of rs475471G>A, rs475533C>T, rs475621A>G and rs475647C>A were all 1, which were completely linked. The results are shown in Fig 3 and Table 1.

### Table 1 Analysis of pairwise linkage disequilibrium base on 9 SNPs of ST3GAL4 gene

<table>
<thead>
<tr>
<th>SNP</th>
<th>475435C&gt;T</th>
<th>475460C&gt;T</th>
<th>475471G&gt;A</th>
<th>475533C&gt;T</th>
<th>475621A&gt;G</th>
<th>475647C&gt;A</th>
<th>475695C&gt;T</th>
<th>475744A&gt;G</th>
</tr>
</thead>
<tbody>
<tr>
<td>475435C&gt;T</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.782</td>
<td>1</td>
</tr>
<tr>
<td>475460C&gt;T</td>
<td>0.185</td>
<td>-</td>
<td>0.995</td>
<td>0.995</td>
<td>0.995</td>
<td>0.995</td>
<td>0.977</td>
<td>1</td>
</tr>
<tr>
<td>475471G&gt;A</td>
<td>0.172</td>
<td>0.921</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>475533C&gt;T</td>
<td>0.172</td>
<td>0.921</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>475621A&gt;G</td>
<td>0.172</td>
<td>0.921</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>475647C&gt;A</td>
<td>0.174</td>
<td>0.921</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>475695C&gt;T</td>
<td>0.175</td>
<td>0.903</td>
<td>0.983</td>
<td>0.983</td>
<td>0.983</td>
<td>0.983</td>
<td>-</td>
<td>0.991</td>
</tr>
<tr>
<td>475743A&gt;G</td>
<td>0.083</td>
<td>0.464</td>
<td>0.431</td>
<td>0.431</td>
<td>0.431</td>
<td>0.431</td>
<td>0.431</td>
<td>-</td>
</tr>
<tr>
<td>475744A&gt;G</td>
<td>0.189</td>
<td>0.979</td>
<td>0.91</td>
<td>0.91</td>
<td>0.91</td>
<td>0.9</td>
<td>0.45</td>
<td>-</td>
</tr>
</tbody>
</table>

Note: The lower triangle is r², the upper triangle is D'.

### Association analysis of genetic variation loci of ST3GAL4 gene with production traits and blood biochemical indexes of Hubbard broilers

The blood samples of 468 Hubbard broilers were sequenced by PCR. After statistical typing, 9 SNPs were associated with production traits of Hubbard broilers, all the 9 SNPs of ST3GAL4 gene were significantly related to ALP (P<0.01). The results were shown in Table 2. In addition, rs475471G>A, rs475533C>T, rs475621A>G and rs475647C>A were significantly associated with body weight and tibial length (P<0.05).

### Table 2 Association analysis of 9 SNPs of ST3GAL4 with ALP of Hubbard broilers
Association analysis of ST3GAL4 genotype and economic traits of F$_2$ resource population

In order to identify the association between the mutations of ST3GAL4 gene and production traits in chicken, especially with ALP. The rs475471G>A of four complete linkage sites was selected as tag SNP, and Hinfl restriction enzyme was used for natural enzyme digestion, which enzyme products were electrophoretic with 2% agarose gel. The typing map was shown in Fig 4, and different genotypes were identified according to agarose gel electrophoresis. The amplification product of GG genotype was a single band with the size of 589 bp, the amplification product of GA genotype was three bands with the
size of 589bp, 451bp and 138bp, respectively, the amplification product of AA genotype was two bands with the size of 451 bp and 138 bp. Because of the 138 bp band was too weak, only the other two bands could be used for genotyping.

The results showed that there was a significant correlation between different genotypes and pH value of breast muscle and leg muscle ($P<0.05$). The pH value of chest muscle and leg muscle of AA genotype individuals were lower than that of GG genotype and GA genotype individuals. The different genotypes of the locus were significantly correlated with 4-week body weight, chest width, body slanting length, pelvis breadth and 12 week tibial circumference ($P<0.05$), the AA genotype and GA genotype had significantly higher percentage of full evisceration, half evisceration, head weight and leg weight than GG genotype ($P<0.05$), the most important result of this work was to find a significant differences in ALP levels among different genotypes ($P<0.01$). The results were shown in Table 3.

To analysis of the above results, the GG genotype had the lowest ALP and grew slowly. The ALP level of AA and AG genotypes were significantly higher than that of GG genotype, and exhibit better growth performance than those individuals with GG genotype. This is consistent with the analysis in Hubbard broilers. In summary, allele A is positively correlated with ALP content, and the genotype with high alkaline phosphatase content was the dominant genotype.

**Table 3** Association analysis results of 475471G>A polymorphic loci of *ST3GAL4* gene and economic traits of F$_2$ resource population

<table>
<thead>
<tr>
<th>Traits</th>
<th>GG=238</th>
<th>GA=407</th>
<th>AA=94</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPH</td>
<td>6.10±0.02c</td>
<td>6.09±0.01b</td>
<td>5.97±0.03a</td>
<td>0.002</td>
</tr>
<tr>
<td>LPH</td>
<td>6.65±0.01c</td>
<td>6.67±0.01b</td>
<td>6.59±0.03a</td>
<td>0.044</td>
</tr>
<tr>
<td>BW4 (g)</td>
<td>314.35±6.93b</td>
<td>324.20±6.66b</td>
<td>326.31±7.87a</td>
<td>0.014</td>
</tr>
<tr>
<td>CW4 (cm)</td>
<td>4.03±0.03b</td>
<td>4.13±0.03a</td>
<td>4.06±0.05ab</td>
<td>0.025</td>
</tr>
<tr>
<td>BSL4 (cm)</td>
<td>11.30±0.09b</td>
<td>11.41±0.08ab</td>
<td>11.52±0.11a</td>
<td>0.043</td>
</tr>
<tr>
<td>PW4 (cm)</td>
<td>5.09±0.03b</td>
<td>5.18±0.03a</td>
<td>5.17±0.05a</td>
<td>0.027</td>
</tr>
<tr>
<td>SEW (g)</td>
<td>81.06±0.23b</td>
<td>81.53±0.21a</td>
<td>81.63±0.28a</td>
<td>0.011</td>
</tr>
<tr>
<td>EW (g)</td>
<td>67.71±0.33b</td>
<td>68.03±0.32ab</td>
<td>68.31±0.36a</td>
<td>0.026</td>
</tr>
<tr>
<td>LW (g)</td>
<td>146.21±3.65b</td>
<td>150.78±3.52a</td>
<td>152.83±4.13a</td>
<td>0.024</td>
</tr>
<tr>
<td>ALP(U/mL)</td>
<td>592.56±57.83c</td>
<td>776.94±54.23b</td>
<td>886.46±70.33a</td>
<td>2.65E-07</td>
</tr>
</tbody>
</table>
CPH, Chest muscle PH, LPH, leg muscle PHBW4, Body Weight at 4 weeks CW4, Chest width at 4 weeks, BSL4, Body slant length at 4 weeks, PW4, Pelvic width, SEW, Semi-evisceration weight, EW, Evisceration weight, LW, Leg weight, ALP, Alkaline phosphatase.

**Genotypic and allelic frequencies and related genetic parameters for the chicken *ST3GAL4* gene**

To detect the variations among commercial broilers, egg-type chickens, dual-type chickens, we performed a PCR analysis on 4825 individuals from 18 populations were analyzed, and to investigate the genotype distribution and gene frequencies of *ST3GAL4* in different populations in this study. The number of genotypes of each variety is shown in Table S2. As shown in Table 4, the genotype frequencies of CB, R308, B817, HBD broilers and F₂ resource groups were significantly different from those of egg-type chickens and dual-type chickens. GG genotype mainly existed in egg-type chickens and dual-type chickens, while AA genotype mainly distributed in commercial broilers and F₂ resource population.

PIC, He and Ne are not only indicators to measure the degree of allele polymorphism and gene mutation, but also indicators to reflect the genetic variation within the population. According to the standard polymorphism information, PIC < 0.25 represents low polymorphism, 0.25 < PIC <0.50 represents moderate polymorphism, while PIC >0.50 represents high polymorphism. The F₂ resource population and commercial broilers all showed moderate polymorphism, which indicated that the F₂ generation and commercial broilers exhibited greater genetic variation and selection potential than egg-type chickens and dual-type chickens. Allele A is positively correlated with ALP and mainly existed in commercial broilers and F₂ generation, what's interesting is that the growth performance of individuals with high ALP level is better, which might be due to the result of artificial selection.

**Table 4** Genotypic and allelic frequencies and related genetic parameters for the chicken *ST3GAL4* gene
<table>
<thead>
<tr>
<th>Breeds/n</th>
<th>Genotypic and allelic frequencies</th>
<th>He</th>
<th>Ne</th>
<th>PIC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GG</td>
<td>GA</td>
<td>AA</td>
<td>G</td>
</tr>
<tr>
<td>AA/597</td>
<td>0.035</td>
<td>0.628</td>
<td>0.337</td>
<td>0.349</td>
</tr>
<tr>
<td>CB/224</td>
<td>0.259</td>
<td>0.598</td>
<td>0.143</td>
<td>0.558</td>
</tr>
<tr>
<td>R308/179</td>
<td>0.274</td>
<td>0.570</td>
<td>0.156</td>
<td>0.559</td>
</tr>
<tr>
<td>B817/71</td>
<td>0.141</td>
<td>0.704</td>
<td>0.155</td>
<td>0.493</td>
</tr>
<tr>
<td>HBD/468</td>
<td>0.297</td>
<td>0.538</td>
<td>0.165</td>
<td>0.566</td>
</tr>
<tr>
<td>F2/739</td>
<td>0.322</td>
<td>0.551</td>
<td>0.127</td>
<td>0.597</td>
</tr>
<tr>
<td>YY/43</td>
<td>0.814</td>
<td>0.186</td>
<td>0.00</td>
<td>0.907</td>
</tr>
<tr>
<td>DX/186</td>
<td>0.758</td>
<td>0.226</td>
<td>0.016</td>
<td>0.871</td>
</tr>
<tr>
<td>GC/78</td>
<td>0.833</td>
<td>0.154</td>
<td>0.013</td>
<td>0.910</td>
</tr>
<tr>
<td>GS/177</td>
<td>0.655</td>
<td>0.316</td>
<td>0.028</td>
<td>0.814</td>
</tr>
<tr>
<td>HD/158</td>
<td>0.842</td>
<td>0.158</td>
<td>0.000</td>
<td>0.921</td>
</tr>
<tr>
<td>HL/307</td>
<td>0.57</td>
<td>0.417</td>
<td>0.013</td>
<td>0.779</td>
</tr>
<tr>
<td>LS/92</td>
<td>0.554</td>
<td>0.413</td>
<td>0.033</td>
<td>0.761</td>
</tr>
<tr>
<td>WH/288</td>
<td>0.729</td>
<td>0.236</td>
<td>0.035</td>
<td>0.847</td>
</tr>
<tr>
<td>XC/384</td>
<td>0.833</td>
<td>0.159</td>
<td>0.008</td>
<td>0.913</td>
</tr>
<tr>
<td>RW/221</td>
<td>0.778</td>
<td>0.213</td>
<td>0.009</td>
<td>0.885</td>
</tr>
<tr>
<td>CS/95</td>
<td>0.663</td>
<td>0.274</td>
<td>0.063</td>
<td>0.800</td>
</tr>
<tr>
<td>FB/113</td>
<td>0.584</td>
<td>0.398</td>
<td>0.018</td>
<td>0.783</td>
</tr>
</tbody>
</table>

AA, Arbor Acres broiler, CB, Cobb broiler, R308, Ross308, B817, 817 broiler, HBD, Hubbard broiler, F2, F2 resource population, YY, Yunyang chicken, DX, Dongxiang chicken, GC, Henan gamecock, GS, Gushi chicken, HD, Houdan chicken, HL, Hy-line brown hen, LS, Lushi chicken, WH, Wuhei chicken, XC, Xichuan chicken, RW, Recessive white chicken, CS, Changshun chicken, FB, Field Broiler. He, expected heterozygosity; Ne, effective allele numbers; PIC, polymorphism information content.

**Expression of different genotypes of ST3GAL4 gene in different tissues**

By analyzing the expression of different genotypes of ST3GAL4 gene in different tissues, we found that ST3GAL4 is most highly expressed in the duodenum, while almost not expressed in the pectoral and leg muscle Fig.5. In addition, the expression level of GG and GA genotype were significantly higher than
that of AA genotype in spleen \((P<0.05)\), the GA and AA genotype were extremely significantly higher than that of GG genotype in duodenum \((P<0.01)\).

**Discussion**

Here, we identified 9 SNPs of the *ST3GAL4* gene located on chromosome 24 for the first time. The variation of SNPs is extremely widespread and can occur at any position in the DNA sequence of the genome, including the coding region and the non-coding region of the gene. A group of interrelated SNPs located in a region on the same chromosome could form haplotypes. More information about the origin structure can be obtained from the distribution of haplotypes. The correlation analysis between haplotypes and actual target phenotypes can effectively find the target genes in the process of heredity and mutation [22]. Because *ST3GAL4* is located within the QTL associated with ALP, we speculated that *ST3GAL4* might also be related to the ALP of chickens, and interestingly, by PCR-SBT of Hubbard broiler population, 9 SNPs that may be linked were identified. The Linkage disequilibrium (LD) refers to the nonrandom association among alleles of different loci. By calculating and analyzing the correlation degree between the existing genetic markers and the phenotypic values of the target trait, the markers significantly associated with the trait can be obtained, which can be used to further screen candidate genes for subsequent studies [23]. Therefore, linkage disequilibrium analysis was carried out for these loci. \(D'\) and \(r^2\) are the measurement values to evaluate the linkage disequilibrium relationship. The size of \(D'\) reflects the frequency of recombination events in the LD region. Some researchers believe that two loci of \(D'>0.8\) are in a state of strong linkage disequilibrium [24, 25]. In addition, the size of \(r^2\) is closely related to the effectiveness of association analysis [26]. When \(r^2>0.33\), the two markers were closely linked [27]. Therefore, \(r^2>0.33\) and \(D'>0.8\) are defined as the criteria of strong linkage disequilibrium. In genetic research, when several SNPs are in linkage disequilibrium, this mutation may be more valuable and effective than a single locus [28]. Therefore, the sequencing results of 468 Hubbard broilers showed that the \(D'\) and \(r^2\) values of rs475471G > A, rs475533C > T, rs475621A > G and rs475647C > A were all 1, which were completely linked, and it was significantly correlated with ALP in this study.

ALP is a marker of osteoblast maturation and an important indicator of bone metabolism. The change of ALP can reflect the bone metabolism of animals to a certain extent [29]. Our previous study found that the serum ALP content of VVD broilers was significantly lower than that of healthy broilers [30]. In order to confirm that *ST3GAL4* is the major gene associated with ALP, rs475471G > A of the above four fully linked loci was selected as the tag SNP, and the appropriate endonuclease (Hinfl) was selected for enzyme typing in \(F_2\) resource population. The results are still highly correlated with ALP. The detection of polymorphic loci related to rs475471G > A in chickens of different breeds showed that the genotypes were different between broilers and dual-type chickens. The AA genotype ALP level was the highest and GG genotype was the lowest. Interestingly, genotypes with high ALP showed better growth performance, and the frequency of dominant genotype and dominant allele was higher in broilers. Besides, the results of tissue expression showed that the expression level of GA and AA genotype in duodenum were significantly higher than that of GG genotype. Duodenum has important digestive function, studies have
found that it plays an important role in functional dyspepsia [31]. This may be due to the low ALP level and poor digestive function of GG genotype in the process of commercial breeding of broilers, which would result in the gradual selection of genotype AA during this process, and thus, genotype AA was more common in broiler population. Furthermore, we speculate that \( ST3GAL4 \) is the major gene regulating ALP.

\( ST3GAL4 \) is the major \( \alpha_2,3 \)-sialyltransferase that acts on N-glycans, the structures of Glycan are diverse and dynamically change to enable complex multicellular systems, such as immunity, the nervous system and development [32–34]. Furthermore, undesirable changes in glycan structures can cause various diseases, including cancer and diabetes [35–37]. A novel SNP in the \( ST3GAL4 \) gene was identified that associated with Von Willebrand factor (VWF) levels and protects coagulation factor VIII (FVIII) activity [38]. And a SNP at rs11220462 of \( ST3GAL4 \) gene was found that associated with cholesterol (T-CHO) and low density lipoprotein concentrations (LDL-C) in human [19]. But variations in the chicken \( ST3GAL4 \) gene remain undefined. Our previous GWAS found a SNP at \( ST3GAL4 \) was significantly correlated with ALP level in chicken (unpublished data), the ALP plays an important role in bone growth and development, it can reflect the activity of bone metabolism and participate in the process of bone calcification [6, 7], and plays an important role in the pathogenesis of OA [15]. In this study, we identified \( ST3GAL4 \) as a major gene associated with ALP. Therefore, \( ST3GAL4 \) gene may affect the normal growth and development of bone by regulating ALP level, which will also affect the performance of chickens to some extent.

**Conclusions**

Collectively, we first time determined 9 SNPs within \( ST3GAL4 \) were related to ALP in chickens, 4 of them were complete linkage sequence variants and the mutant individuals exhibited the dominant genotype. Through the analysis of the association results of Hubbard population and \( F_2 \) resource population, we determined that \( ST3GAL4 \) is a major gene related to ALP. In conclusion, our study first time to report the causal variant of chicken serum ALP level located at chromosome 24, which can predict the serum ALP level of chicken early and effectively, and can be used as a molecular marker for genetic improvement of chicken breeds. This study might also provide a reference for future studies of molecular mechanism about \( ST3GAL4 \) gene.

**Methods**

**Laboratory animals and data collection**

The Hubbard broilers population was described by Guo et al [20] Tang et al [39]. In detail, Hubbard broiler breeds for the commodity generation were reared at a commercial farm in the same environment and feeding management, free access to food and water. The Gushi-Anka chicken \( F_2 \) generation resource population was described by Han et al [40]. To produce the \( F_2 \) resource population, two hatchings were obtained from an \( F_1 \) generation constructed via reciprocal crossing between Chinese native Gushi
chickens and Anka broilers (4 Anka × 24 Gushi, orthogonal, 2 Gushi × 12 Anka, reverse cross, 70 F₁ individuals). The F₂ generation was offspring by mating with hens from other families and consisted of 7 families, which Anka broilers were the male parents of four orthogonal lines, Gushi chickens represented the male parents of the three reverse cross, which comprised 42 grandparents, 70 F₁ parents, and 808 F₂ chickens.

The Hubbard broilers were euthanized with pentobarbitone sodium at 40 days and F₂ generation were euthanized by cervical dislocation followed by decapitation at 84 days. At the time of slaughter, two blood samples were collected from the jugular vein. One sample was placed in an anticoagulant tube to extract DNA and then stored at -20 °C, the other sample was placed in a centrifuge tube to separate the serum and stored at -40 °C. In addition, growth traits of F₂ generation were measured at 0 days and 2, 4, 6, 8, 10 and 12 weeks. After blood collection, 808 individuals were slaughtered to determine the slaughter traits and blood biochemical indexes. A detailed description can be found in previous study [41, 42]. Besides, the Cobb broilers were euthanized at 35 days, which collected the heart, liver, spleen, lung, kidney, pectoral, leg muscle, thymus, duodenum and cartilage to construct the tissue expression profiles.

A total of 4825 DNA samples were collected from 18 populations, including local varieties, layer varieties and broiler varieties. Arbor Acres broiler(AA, n = 597), Ross 308 broiler(R308, n = 179), Hubbard broiler(HBD, n = 468), Cobb broiler(CB, n = 629), 817 broiler(B817, n = 71), Recessive white chicken(RW, n = 221), Gushi-Anka chicken F₂ generation resource population(F₂, n = 739), Hy-line brown hen(HL, n = 307), Dongxiang green-shell layers(DX, n = 186), Xichuan black bone chickens(XC, n = 384), Gushi chickens(GS, n = 177), Changshun green-shell layers(CS, n = 95), Lushi green-shell layers(LS, n = 92), Yunyang chicken(YY, n = 43), Wuhei chicken(WH, n = 288), Houdan chicken(HD, n = 158), Henan gamecock(GC, n = 78), Field broiler(FB, n = 113), respectively. Blood samples of these varieties were obtained from the laboratory for DNA extraction and typing.

Genomic DNA extraction and PCR

Genomic DNA was extracted from blood using DNA extraction kit (TaKaRa, Beijing, China). The primers used in this study were designed by Premier Primer 6.0 software, and synthesized by Sangon Biotech Company (Shanghai, China). The primers used in this study were listed in Table 5.

DNA from each individual of Hubbard broiler was subjected to all SNPs identification and polymerase chain reaction-sequence based typing (PCR-SBT), which was performed by Sangon Biotech Company (Shanghai, China). Linkage disequilibrium analysis of 9 SNPs (rs475435C > T, rs475460C > T, rs475471G > A, rs475533C > T, rs475621A > G, rs475647C > A, rs475695C > T, rs475743A > G, rs475744A > G) showed that 4 SNPs (rs475471G > A, rs475533C > T, rs475621A > G, rs475647C > A) were completely linked. The tag SNP rs475471G > A in ST3GAL4 gene was selected and through online software WatCut (http://watcut.uwaterloo.ca/) to choose the appropriate restriction enzymes Hinfl, which was used for genotyping in F₂ generation resource population and other fifteen varieties population. The restriction primers were consistent with the polymorphic sites detection primers. The PCR amplification was
performed as follows: The preincubation stage (95°C/5 min), 35 cycling stages (95°C/30 s, 63°C/30 s and 72°C/30 s), extension stage (72°C/10 min) and cooling to 4°C.

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Primer Sequence (5’→ 3’)</th>
<th>Product size(bp)</th>
<th>Annealing temperature(℃)</th>
<th>Accession number</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST3GAL4 polymorphic sites detection</td>
<td>F: CCCTTTCCAGTCCCCATTCC R: GGGGTACGTAACCCATTG</td>
<td>589</td>
<td>63</td>
<td>NC_006111.5</td>
</tr>
<tr>
<td>RT-ST3GAL4</td>
<td>F: CGTGGAAGAAGTGTGAGCGG R: CCCCAATGGGACATACGGTC</td>
<td>208</td>
<td>60</td>
<td>XM_417860.6</td>
</tr>
<tr>
<td>GAPDH</td>
<td>F: GAACATCATCCCAGCGTCCA R: CGGCAGGTCAGGTAACAAAC</td>
<td>132</td>
<td>60</td>
<td>NM_204305</td>
</tr>
</tbody>
</table>

RNA extraction, cDNA synthesis and Quantitative PCR analysis

To study the expression of ST3GAL4 in different genotypes, the total RNA extracted from heart, liver, spleen, lung, kidney, pectoral, leg muscle, thymus, duodenum and cartilage from Cobb broiler. The RNA integrity and concentration were estimated using a NanoDrop spectrophotometer (Thermo Fisher Scientific, USA), and verified electrophoretically using an agarose gel. Only qualified samples were used for further analysis. The cDNA synthesis was performed using a PrimeScript RT reagent kit with gDNA Eraser (Takara). qPCR was performed on a LightCycler1 96 real time fluorescent quantitative PCR instrument (Roche Applied Science, Indianapolis, USA), the cycling conditions were as follows: initial denaturation (95 °C/5 min); followed by 30 cycles (95 °C/15 s, 60 °C/30 s, 72 °C/30 s), and finally 10 min extension at 72 °C. The relative quantifications of genes were analyzed using the 2^{-ΔΔCt} method and one-way analysis of variance (ANOVA), respectively [43]. NCBI Primer-Blast was used to design the primers with chicken GAPDH as an endogenous control [44]. The primers used in this study were listed in Table 5.

The significance of the data was analyzed by IBM SPSS (SPSS for Windows, Standard version 24; SPSS, USA). Data obtained was presented as means ± SEM, \( P \leq 0.01 \) indicated an extremely significant difference, \( 0.01 \leq P \leq 0.05 \) indicated a significant difference and \( P > 0.05 \) indicated no difference, and the figure was drawn using GraphPad Prism 6 (GraphPad Software Inc., 2007, San Diego, CA, USA).

**Statistical analysis**

Association study was accomplished with mixed linear model using SPSS 24.0 between genotype and traits of case – control Hubbard broiler population. The statistical model as follows:

\[
y = X\beta + Yb + Zg + e
\]
where $Y$ is the vector of phenotypes, $\beta$ is the Fixed effect vector including to gender (male; female), chicken house (1; 2) and herd (Normal, VVD), $b$ is the additive effects of test SNPs, $e$ is a vector of residuals.

Another association study between the genotypes and selected traits of the Gushi×Anka F$_2$ chickens were performed using SPSS 24.0 according to the following two linear mixed models.

Model I was used to evaluate growth traits. Considering the effects of body weight on carcass traits, carcass weight was used as a concomitant variable in model II, which was applied to calculate carcass traits.

Model I: $Y_{ijklm} = \mu + G_i + S_j + H_k + f_l + e_{ijklm}$

Model II: $Y_{ijklm} = \mu + G_i + S_j + H_k + f_l + b (W_{ijklm} - W) + e_{ijklm}$

Where $Y_{ijklm}$ is the observed value; $\mu$ represents the overall population mean; $G_i$ is the fixed effect of the genotype ($i = GG, GA, AA$); $f_l$ is the fixed effect of the family ($l = 7$); $S_j$ is the fixed effect of sex ($j = 2$); $H_k$ is the fixed effect of the hatch ($k = 2$); $b$ represents the regression coefficient for carcass weight; $W$ is the average slaughter weight; $W_{ijklm}$ is individual slaughter weight; and $e_{ijklm}$ is the random error. $G_i, S_j, H_k$ are fixed factors, and $f_l$ is a random factor.

Above multiple comparisons were analyzed with the least square means to display significant differences among genotypes with Bonferroni’s test.

SPSS 24.0 was used to analyze the correlation between genotype and traits of Hubbard broiler and F$_2$ generation resource population. Bonferroni’s multiple comparison method was performed for differences among genotypes[45]. All Data obtained by statistical analysis was presented as means ± SEM. $P \leq 0.01$ indicated an extremely significant difference, $0.01 \leq P \leq 0.05$ indicated a significant difference and $P > 0.05$ indicated no difference.

Haplotypes were constructed on the basis of the nine SNPs in case – control Hubbard broiler population by use of the Haploview program. The function of this program is to reconstruct haplotypes from population data.

**Abbreviations**

QTLs: quantitative trait loci; ALP: Alkaline phosphatase; GWAS: genome-wide association studies; BW: body weight; CW: chest width; BSL: body slanting length; PW: pelvis width; SEW: semi-evisceration weight; EW: evisceration weight; LW: Leg weight; CPH: Chest muscle PH; LPH: leg muscle PH; OA: osteoarthritis; ECM: extracellular mechanism; T-CHO: cholesterol; LDL-C: low density lipoprotein concentrations; VVD: Valgus-varus deformity; LD: linkage disequilibrium; AA: Arbor Acres broiler; CB: Cobb broiler; R308: Ross308; B817: 817 broiler; HBD: Hubbard broiler; F$_2$: F$_2$ resource population; YY: Yunyang chicken; DX:
Dongxiang chicken; GC: Henan gamecock; GS: Gushi chicken; HD: Houdan chicken; HL: Hy-line brown hen; LS: Lushi chicken; WH: Wuhei chicken; XC: Xichuan chicken; RW: Recessive white chicken; CS: Changshun chicken; FB: Field Broiler; He: expected heterozygosity; Ne: effective allele numbers; PIC: polymorphism information content; VWF: Von Willebrand factor. FVIII protects coagulation factor VIII.

**Declarations**

**Supplementary Information**

Table S1. Typing results of Hubbard broiler by PCR-SBT

Table S2. Number of genotypes of each variety

**Acknowledgments**

The authors gratefully acknowledge all professors and students in the Henan Key laboratory for innovation and utilization of chicken germplasm resources.

**Authors’ Contributions**

H.T. performed the most experiment and wrote manuscript. Y.M., J.L., Z.Z., C.C., and L.Z. assisted with sample collection and perform the tissue expression experiment. W.L., H.H. contributed to data curation and validation. Z.L., R.J., Y.T., G.L., F.Y., G.S., and Y.W. contributed to sample collection and construction of F₂ resource population. D. L., Y. Z., H. L., X., L. helped perform the analysis with constructive discussions. R.H. designed the study and revised manuscript., X.K. conceived the study and project administration.

**Funding**

This work is supported by the Scientific Studio of Zhongyuan Scholars (30601985), the Key Scientific Research Projects of Henan Colleges and Universities (20A230006), the Innovation Research Team of the Ministry of Education (IRT16R23).

**Availability of data and materials**

The data are shown in the main manuscript and supplemental materials.

**Ethics statement**

All the animal experiments were performed according to the Regulations of the Chinese National Research Council (1994) and approved by the Animal Care Committee of the College of Animal Science and Technology, Henan Agricultural University, China (Permit Number: 18–0085). All experimental procedures and methods were carried out in accordance with approved guidelines to ensured animal welfare.
Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

1 College of animal science and technology, Henan Agricultural University, Zhengzhou, Henan Province 4500046, China

2 Henan Key laboratory for innovation and utilization of chicken germplasm resources, Zhengzhou, Henan Province 450046, China

References


**Figures**
Figure 1

Sequencing maps of heterozygous mutations of 9 SNPs in the ST3GAL4 gene.

ST3GAL4
Chromosome 24: 473593-482511

Exon1 Exon2 Exon3 Exon4 Exon5 Exon6 Exon7 Exon8 Exon9 Exon10 Exon11 Exon12 Exon13
5'UTR 23 137 Exon1 76 Exon4 79 70 81 98 Exon5 Exon6 Exon7 61 Exon8 96 190 Exon9 Exon10 144 Exon11 Exon12 Exon13 144 87 3'UTR

Exon Intron
Figure 2

The position of 9 SNPs on the ST3GAL4 gene

A

B

C

Figure 3

The block of linkage disequilibrium analysis based on D’

The block of linkage disequilibrium analysis based on r2. The darker the color, the greater of the D’ and r2.

Haplotype

Figure 4

bp

2000

1000

750

500

250

100

589

451

138
Figure 5

Different genotypes of ST3GAL4 gene expressed in different tissues, *P < 0.05, **P < 0.01.

Supplementary Files

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

- TableS1.xlsx
- TableS2.xlsx