The polymorphisms in ATPase sarco / endoplasmic reticulum Ca2+ transporting-3 gene associated with gastrointestinal nematode infection in goat

Ali Akbar Bhuiyan[1,2], Pan Ni[1], Abdelmotaleb Ahmed Elokil[1,3], Jingjin Li[1], Anjuman Ara Bhuyan[4], Shuhong Zhao[1] and Xiaoyong Du[1,5]

[1]Key Laboratory of Agricultural Animal Genetics, Breeding and Reproduction, Ministry of Education, College of Animal Science and Veterinary Medicine, Huazhong Agricultural University, Wuhan 430070, P. R. China.
***(Corresponding author:** Ali Akbar Bhuiyan (A. A. Bhuiyan): aab76_blri@yahoo.com; Cell: +86-15927465507)*
[2]Senior Scientific Officer, Bangladesh Livestock Research Institute, Savar, Dhaka-1341, Bangladesh
[3]Department of Animal Production, Faculty of Agriculture at Moshtohor, Benha University, Egypt
[4]State key Laboratory of Agricultural Microbiology, Ministry of Education, College of Animal Science and Veterinary Medicine, Huazhong Agricultural University, Wuhan 430070, P. R. China.
[5]Hubei Key Laboratory of Agricultural Bioinformatics, College of Informatics, Huazhong Agricultural University, Wuhan 430070, P. R. China.

Summary

Gastrointestinal nematode (GIN) is the brutal pathogenic parasite in small ruminants. The aim of our study is to identify single nucleotide polymorphisms in the ATP2A3 gene and investigate their association with GIN infection traits measured in 384 goats encompassing three Chinese indigenous goat populations. The seven novel variants were identified in the goat ATP2A3 gene, in where mutations located in 3′ UTR and two synonymous mutations located in exon regions. Genotypic and allelic frequencies of seven variants were consistent with Hardy-Weinberg equilibrium and the population genetics parameters Association analysis revealed that out of seven polymorphisms, four polymorphisms had highly significant effects on GIN infection traits in goat. From haplotype and LD structures of seven SNPs, two pair variants, C24361893T and C24379974T from exon region as well as C24358400A and G24358441C from 3′ UTR region showed significant LD with each other with strong correlations and spanning 18 kb and 0 kb blocks, respectively. From four polymorphic sites of ATP2A3 gene and their haplotype combinations, eight previously undiscovered haplotypes were constructed from these four mutations, where the AG and CC haplotypes showed the highest frequency 65.0% and 53.3%, followed by others haplotypes. The qRT-PCR analysis, the relative expression of mRNA in ATP2A3 gene was higher in susceptible goats compared to that of resistant goats. Our results indicated that the ATP2A3 gene is a good possible candidate gene either as a major gene or as an associated with a major gene in GIN infection traits in goat.

Key words: SNP; Genetic resistant; ATP2A3; GIN infection; Goat

Introduction

Gastrointestinal nematode (GIN) infection is one of the most annoying parasitic diseases for livestock producers because of its negative effects on the production performance, ultimately reduces economic return (Bishop, 2012; Bressani et al., 2014). The haematophagous GIN, especially *Haemonchus contortus* is the brutal pathogenic parasite in subtropical and tropical countries (Prasad et al., 2016). Each adult Haemonchus can suck 0.05 to 0.07 ml blood from host per day (Malviya et al., 1979). This blood sucking GIN infects the gastrointestinal tract of ruminants. In addition, this parasite, especially *Haemonchus contortus* is overwhelmed part in the ruminant livestock farming, because of its anthelmintic drugs resistance (Howell et al., 2008). It becomes resistant due to frequently use of anthelmintic drugs without the definitive diagnosis, as a result control of infection is becoming more difficult and also increasing treatment cost to the ruminant farm.

The recent advance in molecular genetics has provided an opportunity to identify different genes and pathways associated with different production/ or economic traits and diseases related traits. The identification of candidate genes that influence the individual's production performance or diseases control would provide a better understanding of the physiological processes, like parasites susceptibility or resistant. Improving suitable traits by using traditional breeding programs would be difficult as such traits are intricate and under the control of many genes. A better understanding of the genetic basis of these traits is desired to efficiently escort the targeted genetic upgrading of animal species (Sun et al., 2013) through marker-assisted selection (Sun et al., 2013). Therefore, selection of genetic marker has been taking placed on priority for the rapid improvement of molecular genetics (Dodgson et al., 1997; Vignal et al., 2002). One important marker-assisted selection is single nucleotide polymorphisms (SNPs), which are usually used to complex traits mapping by linkage disequilibrium (LD) when they are associated with the interested gene (Emara and Kim, 2003). Previous several studies have identified a number of polymorphisms in various genes that are associated with the traits of GINs in goat. Toll-like receptors (TLRs), Beta Tubulin Isotype-1, Dopamine receptor binding-1 (DRB-1), Insulin-like growth factor-1, Interferon gamma (IFNG), Interleukins (ILs) genes have been implicated for the FEC trait in goat (Alim et al., 2016; Asif et al., 2016).

We also explored genes related to GIN infection in our previous RNA sequencing study that were closely associated with susceptible and resistant goat (Bhuiyan et al., 2017). One important differentially expressed gene (DEG) namely the sarco/ endoplasmic reticulum calcium-ATPase3 gene (ATP2A3) can be a functional candidate gene showing 4.05867 fold change (log2) up-regulated expression in susceptible goat groups were chosen for the identification marker-assisted selection, which belong to the Sarco/ Endoplasmic Reticulum calcium ATPases (SERCAs) family. SERCAs are a protein family possess multiple isoforms, where ATP2A3 genes encoding SERCA3 isoform (Floresperedo et al., 2016). No polymorphisms in the ATP2A3 gene that may affect any disease traits in goats have been investigated. Therefore, it is the first study to identify SNPs in mRNA sequence of ATP2A3 gene that describes GIN infection associations with the goat. This will provide further useful and detailed information that can be used to advance goat breeding by molecular marker-assisted selection programs.

Materials and methods

Animal selection and generation of phenotypic data as a GIN infection traits

A total of 384 goats were selected from three different locations in southern China including three local goat populations named as Enshi black (ESB) goat, Nanjiang yellow (NJY) goat, and Yichang white (YCW) goat. Animals were monitored for parasitological (FEC) and hematological parameters (Hgb and PCV). FEC was determined by using modified McMaster technique (Roepstorff and Nansen, 1998) and blood parameters were determined using Mindray (BC-2800Vet) Auto Hematology Analyzer (Shenzhen Mindray Bio-Medical Electronics Co. Ltd, Shenzhen, China).

Challenge trial, extreme animal selection for RNA sequencing and gene selection for polymorphisms detection

The challenge trials by *Haemonchus contortus* were performed at *Lao Gao Huang* goat farm, Yichang, Wuhan, China. Considering the phenotypic data and the challenge trial performance, 4 resistant and 4 susceptible goats were selected for RNA sequence. We analyzed RNA sequences and found some up and down regulated DEGs (data not presented here) related to GIN infection traits and ATP2A3 gene showed 4.05867 (log2) fold change up-regulated expression in susceptible goat groups. We also validated this expression using RNA from our resource resistant and susceptible goat populations and got the same result. Considering higher expression values of this gene, we were chosen its mRNA (XM_018064437.1) for further detection polymorphisms as marker-assisted selection. The detail procedures of nematode challenge trial, RNA sequencing and bioinformatics analysis were described in our previous research article (Bhuiyan et al., 2017).

DNA extraction, polymorphisms detection and genotyping

The genomic DNA was extracted from ear tissues of each goats using genomic DNA kit (TianGen, Beijing, China) following the manufacturer's guidelines. The DNA concentration and quality were determined by NanoDrop 2000 (Thermo Fisher Scientific, Waltham, MA, USA) spectrophotometer and gel electrophoresis.

We used goat re-sequencing data to detect possible SNPs in ATP2A3 gene. Caprine mRNA of ATP2A3 gene (xm_018064437.1) was downloaded from NCBI and was setup as a "reference genome" using Burrows-Wheeler Alignment (BWA) index (Li and Durbin, 2009). To obtain more sequence variants, we downloaded published re-sequencing data of 33 goat individuals (SRA accession numbers: ERR219543, ERR219544 ERR219545, ERR297229, ERR313197, ERR313199, ERR313200, ERR313206, ERR313210 ,ERR313213, ERR318225, ERR318226, ERR318227, ERR318228, ERR318229 ERR340337, ERR345980, ERR345981, ERR345982, ERR470101, ERR470102, ERR470103 ,ERR470105, SRR1265081, SRR1265097, SRR1265903, SRR1265904 SRR1265905, SRR1265906, SRR1265907, SRR1265908, SRR1265909, SRS1274086) from diverse populations including breeds from Morocco, France, Iran, Korea and American (Benjelloun et al., 2015; Bickhart et al., 2017). After quality filtering, the sequence reads of each animal was separately mapped to "reference genome" ATP2A3 gene using BWA version 0.5.9 (Li and Durbin, 2009) default parameters BAM files for all animals were sorted, and then were suffer duplicate filtration using Picard version 1.108

(http://broadinstitute.github.io/picard/). Re-aligner Target Creator and Indel Re-aligner were applied for local re-alignment, and individual sample SNP calling were performed using Haplotype Caller (parameter: -pairHMM VECTOR_LOGLESS_CACHING--emitRef Confidence GVCF--variant_ index_type LINEAR--ariant_index_parameter 128000) from the GATK version 3.50 (Mckenna et al., 2010). Multi-sample SNP calling were performed to merge the GVCFs using Genotype GVCFs using default setting from the GATK version 3.50. The identified SNPs in VCF results were filtered in the downstream analysis by requiring a minimum coverage depth of 5, a minimum RMS (root mean square) mapping quality score of 20. We also filtered the SNPs with MAF (minor allele frequency) < 0.05 and with missing genotype > 10% of animals in the population. Much stricter, A Hardy–Weinberg Equilibrium test (HWE) were performed using VCFTOOLs (parameter: --hardy) to observe the frequency of genotypes. A SNP, which present both two homozygous and one heterozygous (HOM1/HET/HOM2) in the population were kept, otherwise it were removed.

The identified SNPs were then genotyped in the 384 goats individuals using matrix-assisted laser desorption ionization time of flight mass spectrometry (MALDI-TOF-MS) assay (Sequenom Mass ARRAY®, Bioyong Technologies Inc. HK).

RNA extraction, cDNA production, and qRT-PCR

The RNA was extracted from peripheral blood of selected extreme animals using the TRIzol reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's protocol to know the relative expression in mRNA of ATP2A3 gene through qRT-PCR. A standard prime scipt™ RT reagent kit with gDNA eraser (perfect real time), TAKARA bio inc. for cDNA production, and SYBR green real time PCR master mix (Toyobo Co., Ltd., Osaka, Japan) for the qRT-PCR reaction was used in this experiment. The RNA and cDNA quality and concentration were detected by Nano Drop (ND 2000, Thermo Fisher Scientific, Waltham, MA, USA) spectrophotometer and gel electrophoresis. The caprine β-actin (housekeeping) gene was used to know the relative expression as an internal control gene for normalization of cDNA loading differences in this experiment. The primers were designed using NCBI primer blast and checked by OLIGO-7 (Molecular Biology Insights) software in accordance with the sequences of the corresponding goat mRNAs in GenBank.

Statistical analysis

The logarithm transformation mathematical techniques were applied before analysis data since the FEC data were not normally distributed, they were positively skewed. The data were transformed into log10 (n+100), where n is the actual value (Chauhan et al., 2003). The value of hemoglobin (Hgb) and pack cell volume (PCV) of each goat were also transformed into log10 (n+100) before analysis. The SeqMan II version 5.01 (DNAStar Inc., Madison, WI, USA) was used to assembled and analyzed the DNA sequences. The sequences of nucleotide and amino acids were analyzed using MEGA-5 software (Tamura et al., 2014). The genotypic and allelic frequencies among all breeds were calculated by the standard procedure (Frankham and Richard, 1996). The goodness of Hardy-Weinberg Equilibrium (HWE) was tested by chi square (χ^2) for different locus-population combinations and the number of observed and effective alleles based on the likelihood ratio test. The indexes of population genetics including polymorphism information contents (PIC), effective allele numbers (Ne) and gene heterozygosity (He) were

calculated according to the Botstein's methods (Botstein et al., 1980) and Nei's methods (Nei and Roychoudhury, 1974).

The association analysis of genotypes and GIN infection traits were analyzed using general linear model (GLM) for least square means (LSMs) with statistical analysis software (SAS) version 9.1.3 (SAS Institute Inc., Cary, NC, USA) including Duncan's post hoc test to separate the means as follows below:

$$Y_{ijkl} = \mu + G_i + B_j + S_k + G_iB_j + G_iS_k + B_jS_j + G_iB_jS_k + e_{ijkl} \quad \text{(I)}$$

Where,
Y_{ijkl} = the trait measured in goats;
μ = the population mean;
G_i = the fixed factor of genotype (i^{th} SNP, i= CC, CT, TT etc.);
B_j = the fixed factor of breed (j^{th} breed, j= ESB, NJY and YCW);
S_k = the fixed factor of sex (k^{th} sex, k= male and female);
G_iB_j = the interaction effect of genotype and breed (i^{th} SNP and j^{th} breed);
G_iS_k = the interaction effect of genotype and Sex (i^{th} SNP and k^{th} Sex);
B_jS_k = the interaction effect of breed and Sex (j^{th} breed and k^{th} Sex);
$G_iB_jS_k$ = the interaction effect of genotype, breed and Sex (i^{th} SNP, j^{th} breed and k^{th} Sex);
e_{ijkl} = associated random residual error.

The LSM and SEM values were computed for all genotype effects. However, the LSM and SEM value were not computed either for sex or for any interaction effects due to the fixed factor of kth sex and most of the interaction effects didn't have considerable significant effects. Analysis of haplotypes association was done by Haploview software (Barrett et al., 2005). The qRT-PCR amplifications were conducted using an independent set of four biological and three technical replicates per sample. The $2^{-\Delta\Delta CT}$ analysis method was performed to calculate the relative quantification of expression values of resistant and susceptible goats (Livak and Schmittgen, 2001; Livak, 2008). Between qPCR datasets of two groups were also compared using t-test and differences were considered significant when $p < 0.05$.

Results

The protein sequences of the ATP2A3 gene of goat, cattle, sheep, buffalo, horse, cat, human, camel, pig, mouse and chicken were aligned and a phylogenetic tree was constructed using MEGA version 6 with the maximum likelihood method. The maximum likelihood phylogenetic tree showed that the caprine ATP2A3 gene was closely related to ATP2A3 gene of sheep, cattle and buffalo than that of other studied species **(Figure 1)**.

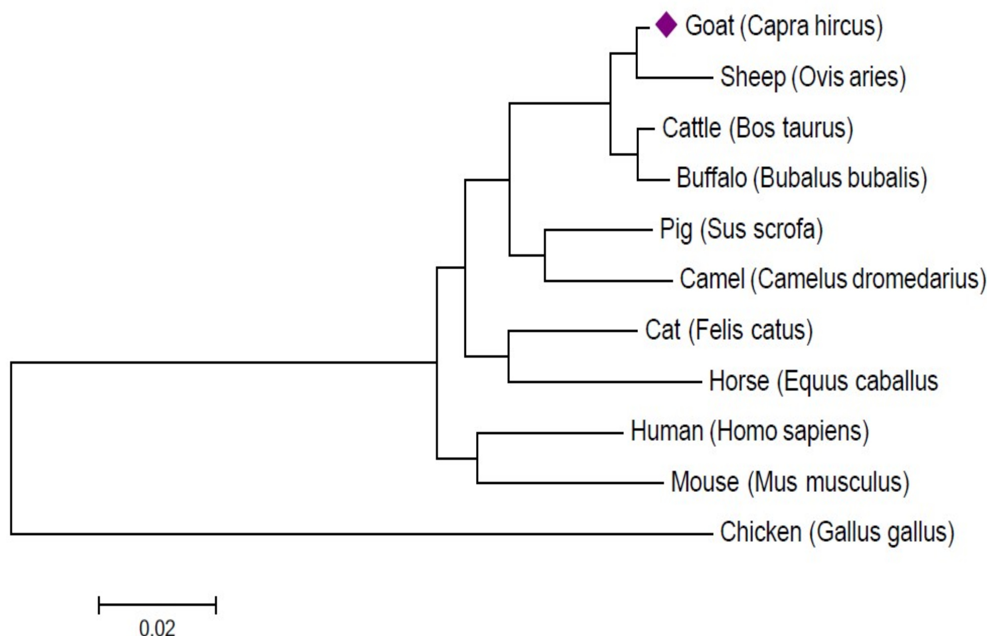

Figure 1: Maximum likelihood phylogenetic tree of ATP2A3 gene in goat

The protein sequences of ATP2A3 gene of goats also showed much closer amino acid similarity to the ATP2A3 gene of sheep (98%) and cattle (98%) than to that of buffalo (97%), pig (96%), Camel (96%), cat (93%), horse (93%), human (93%), mouse (92%) and chicken (80%) (**Table 1**).

Table 1: The protein sequences similarity and values of pair wise comparison of ATP2A3 gene among different species

Species	Goat	Sheep	Cattle	Buffalo	Pig	Camel	Cat	Horse	Human	Mouse	Chicken
Goat	100										
Sheep	98	100									
Cattle	98	97	100								
Buffalo	97	97	99	100							
Pig	96	96	96	96	100						
Camel	96	95	96	96	96	100					
Cat	93	93	93	92	95	95	100				
Horse	93	93	92	92	94	95	94	100			
Human	93	93	93	93	95	95	94	93	100		
Mouse	92	93	92	91	94	95	92	92	93	100	
Chicken	80	80	89	79	82	81	80	80	80	80	100

Pair wise similarity values under triangle as %

Polymorphisms (SNPs) in mRNA sequence of ATP2A3 gene in goat

The seven novel (previously unknown) variants (SNPs) were identified in mRNA sequence of ATP2A3 gene, in where, two were located at exon regions (C24379974T or C408T and C24361893T or C3156T) and five were located in 3′ UTR regions (A24359223G or A3302G, G24358988A, or A3537G, C24358605A or C3920A, G24358441C or G4084C and

C24358400A or C4125A). The information of these seven previously unknown SNPs is presented in **Table 2**. Among the seven SNPs, two were synonymous SNPs at exon region; C24379974T and C24361893T were encoded Glycine and Arginine, respectively, relative to the ATG start codon.

Table 2. Information of seven variants (SNPs) in mRNA sequence of ATP2A3 gene in goat

Sl. no.	SNPs & Locations	SNPs & Positions		Code	Encode**	Amino acid substitution
		In genome	In mRNA*			
1	C/T, (Exon-4)	C24379974T	C408T	GG[C/T]	Gly = Gly	Synonymous
2	C/T, (Exon-22)	C24361893T	C3156T	AA[C/T]	Asn = Asn	Synonymous
3	A/G, (3´-UTR)	A24359223G	A3302G	-	-	-
4	G/A, (3´-UTR)	G24358988A	A3537G	-	-	-
5	C/A, (3´-UTR)	C24358605A	C3920A	-	-	-
6	G/C, (3´-UTR)	G24358441C	G4084C	-	-	-
7	C/A, (3´-UTR)	C24358400A	C4125A	-	-	-

*Reverse complement mRNA sequence of ATP2A3 gene (XM_018064437.1), ** Three letters data-base encodes

Genotypic and allelic frequencies of different SNPs of ATP2A3 gene in goat

Genotypic and allelic frequencies of seven variants in mRNA sequence of ATP2A3 gene were measured. At the C24379974T locus, the homozygous CC genotypes were showed dominant in ESB and YCW goat populations, except NJY goat population. The heterozygous CT genotype was dominant (47.9 %), which was followed by the CC and TT genotypes in the NJY goat populations, and C was identified as an advantageous allele in all populations. But, at the C24361893T locus, the heterozygous CT genotype was dominant in all populations. The frequency of the homozygous GG genotype was highest in the A24359223G mutation in all goat populations, with G being the predominant allele. In the G24358988A and G24358441C polymorphisms, the homozygous GG genotype was also dominant followed by their corresponding genotypes in all goat populations. At the C24358605A locus, the homozygous locus CC genotypes were showed dominant, except ESB population, but at the C24358400A locus, the homozygous locus AA showed dominant, except ESB populations. The ranges of minor allele frequencies (MAF) were 0.085 to 0.445. The chi square ($\chi 2$) analysis showed that the genotypic distribution of three goat populations among different SNPs was consistent with HWE ($P > 0.05$), except for the G24358441C mutation in the YCW goat population, where the CC genotypes were not available. The average PIC, He and Ne values were 0.26 ± 0.042, 0.32 ± 0.061, 1.54 ± 0.133 in ESB; 0.28 ± 0.037, 0.35 ± 0.054, 1.60 ± 0.125 in NJY; and 0.28 ± 0.029, 0.34 ± 0.042, 1.55 ± 0.096 in YCW goat populations, respectively. The lowest PIC, He and Ne values were shown in ESB goat populations compare to NJY and YCW goat populations.

Association with the genotypes of seven polymorphisms in ATP2A3 gene and GIN infection traits of three Chinese goat populations

The descriptive statistics regarding GIN infection traits are presented in **Table 3**. The results indicated that among the three selected traits, the negative correlation was showed between FEC and Hgb ($r = - 0.94$), FEC and PCV ($r = - 0.899$), where as positive correlation

was showed between Hgb and PCV (r = 0.942).

Table 3. Descriptive statistics of GIN infection traits and correlation analysis

Traits (N)	Mean	SEM	Minimum	Maximum	FEC	Hgb	PCV
FEC (384)	2.578	0.217	2.124	3.799	1		
Hgb (384)	2.263	0.012	2.205	2.301	- 0.940**	1	
PCV (384)	2.071	0.007	2.044	2.104	- 0.899**	0.942**	1

FEC = Fecal egg count (egg per gram in feces), Hgb = Hemoglobin (g/L), PCV= Packed cell volume (%), N = Number of individuals, SEM = Standard error of mean, ** = Significant at $P < 0.01$.

The all significant and non significant results with their F-value of association analysis between different variant genotypes of ATP2A3 gene and GIN infection traits of goat are presented in **Table 4**, according to followed model for data analyzing. The results showed significant associations in polymorphisms C24379974T, C24361893T, G24358441C and C24358400A between the variant genotypes and GIN infection traits. The variants C24361893T, G24358441C and C24358400A were also showed significant association between the breed and GIN infection traits. There were no significant differences ($P \geq 0.01$) in sex as well as all of the interaction effects, except interaction effect in G_iS_k and $G_iB_jS_k$ in Hgb trait for variant C24361893T that were showed significant differences at $P < 0.01$.

At the variants C24379974T and C24361893T, the TT genotypes were found significant higher LSM values ($P < 0.001$) than the CC and CT genotypes with the FEC trait, in contrast, significantly higher LSM values ($P < 0.001$) were found in CC genotype than the CT and TT genotypes with Hgb and PCV traits, respectively **(Table 5)**. There were no significant differences found between the genotypes of variants A24359223G, G24358988A and C24358605A and GIN infection traits in this association analysis. Genotypes of the G24358441C variant had a significant association with the FEC, Hgb and PVC traits. Significantly higher LSM values ($P < 0.001$) were found for the GG genotype than for the GC genotypes for FEC trait, but traits Hgb and PCV showed significantly lower LSM values ($P < 0.01$) than the GC values, where there is no available homozygous CC genotype in this mutation. The polymorphism C24358400A had highly significant ($P < 0.001$) associations with GIN infection traits **(Table 5)** and showed highest F-value of significant **(Table 4)**. In this variant, AA genotypes showed highest LSM values than the CC and CA genotypes.

Table 4: Association analysis between the genotypes of seven polymorphisms in mRNA sequence (XM_018064437.1) of ATP2A3 gene and GIN infection traits of three Chinese goat breeds.

SNPs, genome position, mRNA position (location)	**Traits**	**N**	**F-value and level of significance**						
			G	**B**	**S**	**G × B**	**G × S**	**B × S**	**G × B × S**
C/T,	FEC	384	29.77**	2.79NS	0.10NS	2.84NS	0.21NS	0.50NS	0.81NS
C24379974T,	Hgb	384	34.04**	2.04NS	0.29NS	2.95NS	0.68NS	0.25NS	0.76NS
C408T (Exon-4)	PCV	384	30.58**	3.00NS	0.98NS	1.89NS	0.20NS	0.14NS	0.41NS
C/T,	FEC	384	38.59**	7.69*	0.13NS	1.26NS	3.09NS	0.12NS	2.88NS
C24361893T,	Hgb	384	30.07**	12.03**	0.02NS	0.95NS	5.63*	0.83NS	3.42*
C3156T (Exon-22)	PCV	384	23.89**	12.12**	0.36NS	0.68NS	3.89NS	0.59NS	2.85NS
A/G,	FEC	384	1.80NS	4.71**	0.08NS	2.62NS	0.45NS	0.30NS	0.29NS
A24359223G,	Hgb	384	0.73NS	8.11NS	1.06NS	2.65NS	0.40NS	0.25NS	0.06NS
A3302G (3´-UTR)	PCV	384	1.40NS	0.45NS	1.25NS	0.62NS	1.03NS	0.77NS	0.55NS
G/A,	FEC	384	0.73NS	2.30NS	0.04NS	1.02NS	0.40NS	0.08NS	0.81NS
G24358988A,	Hgb	384	0.47NS	2.64NS	0.09NS	1.64NS	0.49NS	0.38NS	0.72NS
A3537G (3´-UTR)	PCV	384	0.50NS	2.84NS	0.37NS	0.92NS	0.41NS	0.27NS	0.65NS
C/A,	FEC	384	0.68NS	3.09NS	0.02NS	2.28NS	0.74NS	0.49NS	0.97NS
C24358605A,	Hgb	384	0.15NS	5.97*	0.66NS	2.16NS	0.91NS	0.33NS	0.54NS
C3920A (3´-UTR)	PCV	384	0.74NS	6.64*	0.88NS	2.07NS	0.84NS	0.39NS	0.66NS
G/C,	FEC	384	15.09**	13.10**	0.17NS	6.34NS	0.33NS	1.44NS	1.45NS
G24358441C,	Hgb	384	10.91*	12.41**	0.63NS	2.98NS	0.31NS	0.15NS	0.47NS
G4084C (3´-UTR)	PCV	384	10.32*	15.29*	1.82NS	4.42NS	0.96NS	0.22NS	0.59NS
C/A,	FEC	384	82.69**	7.08*	0.42NS	0.43NS	0.37NS	0.28NS	1.82NS
C24358400A,	Hgb	384	65.27**	8.36**	0.00NS	1.36NS	0.45NS	0.11NS	1.56NS
C4125A (3´-UTR)	PCV	384	51.24**	9.21*	0.02NS	1.07NS	0.73NS	0.08NS	1.36NS

SNPs=Single nucleotide polymorphisms, FEC=Fecal egg count (egg per gram in feces), Hgb=Hemoglobin (g/L), PCV= Packed cell volume (%),N= Number of individual, G=SNPs genotypes, B = Breeds, S = Sex, G×B = Genotypes and breeds interaction, G×S =Genotypes and sex interaction, B×S=Breeds and sex interaction, G×B×S =Genotypes, breeds and sex interaction,* = Significant at P < 0.01, ** = Significant at P < 0.01, [a,b,c] = LSM values bearing different letters in each row are significantly different at * and **, [NS] = Non significant at P ≥ 0.01.

Table 5: Association analysis between the genotypes of seven polymorphisms in mRNA sequence (XM_018064437.1) of ATP2A3 gene and GIN infection traits of three Chinese goat breeds.

SNPs, genome position, mRNA position (location)	Traits	N	Genotype frequency (LSM ± SEM)			F-value and level of significant
			CC	**CT**	**TT**	
C/T, C24379974T, C408T (Exon-4)	FEC	384	2.40 ± 0.03[c]	2.61 ± 0.03[b]	3.12 ± 0.09[a]	29.77**
	Hgb	384	2.27 ± 0.01[a]	2.25 ± 0.01[b]	2.24 ± 0.03[b]	34.04**
	PCV	384	2.07 ± 0.01[a]	2.06 ± 0.01[b]	2.05 ± 0.01[c]	30.58**
			CC	**CT**	**TT**	
C/T, C24361893T, C3156T (Exon-22)	FEC	384	2.37 ± 0.04[c]	2.58 ± 0.03[b]	2.93 ± 0.04[a]	38.59**
	Hgb	384	2.27 ± 0.01[a]	2.26 ± 0.01[b]	2.25 ± 0.01[b]	30.07**
	PCV	384	2.07 ± 0.01[a]	2.07 ± 0.01[a]	2.06 ± 0.01[b]	23.89**
			AA	**AG**	**GG**	
A/G, A24359223G, A3302G (3´-UTR)	FEC	384	2.70 ± 0.07	2.59 ± 0.04	2.55 ± 0.03	1.80[NS]
	Hgb	384	2.26 ± 0.0	2.26 ± 0.01	2.26 ± 0.01	0.73[NS]
	PCV	384	2.07 ± 0.0	2.07 ± 0.01	2.07 ± 0.01	1.40[NS]
			GG	**GA**	**AA**	
G/A, G24358988A, A3537G (3´-UTR)	FEC	384	2.59 ± 0.03	2.56 ± 0.05	2.73 ± 0.12	0.73[NS]
	Hgb	384	2.26 ± 0.0	2.26 ± 0.01	2.25 ± 0.01	0.47[NS]
	PCV	384	2.07 ± 0.01	2.07 ± 0.01	2.02 ± 0.01	0.50[NS]
			CC	**CA**	**AA**	
C/A, C24358605A, C3920A (3´-UTR)	FEC	384	2.61 ± 0.04	2.55 ± 0.04	2.58 ± 0.06	0.68[NS]
	Hgb	384	2.26 ± 0.01	2.51 ± 0.01	2.36 ± 0.01	0.15[NS]
	PCV	384	2.06 ± 0.01	2.07 ± 0.01	2.05 ± 0.01	0.74[NS]
			GG	**GC**	**CC**	
G/C, G24358441C, G4084C (3´-UTR)	FEC	384	2.55 ± 0.02[b]	2.85 ± 0.07[a]	--	15.09**
	Hgb	384	2.26 ± 0.01[a]	2.25 ± 0.01[b]	--	10.91*
	PCV	384	2.07 ± 0.01[a]	2.06 ± 0.01[b]	--	10.32*
			CC	**CA**	**AA**	
C/A, C24358400A, C4125A (3´-UTR)	FEC	384	3.16 ± 0.05[b]	2.58 ± 0.03[b]	2.36 ± 0.03[a]	82.69**
	Hgb	384	2.24 ± 0.01[c]	2.26 ± 0.01[b]	2.27 ± 0.01[a]	65.27**
	PCV	384	2.05 ± 0.01[b]	2.07 ± 0.01[a]	2.07 ± 0.01[a]	51.24**

LSM = Least squares mean, SEM = Standard error of mean, SNPs=Single nucleotide polymorphisms, FEC=Fecal egg count (egg per gram in feces), Hgb=Hemoglobin (g/L), PCV= Packed cell volume (%),N= Number of individual, G=SNPs genotypes (*i.e.* CC, CT, TT),* = Significant at $P < 0.01$, ** = Significant at $P < 0.001$,[a,b,c] = LSM values bearing different letters in each row are significantly different at * and **, NS = Non significant at $P \geq 0.01$.

Haplotype and linkage disequilibrium (LD) analysis of ATP2A3 gene in goat populations

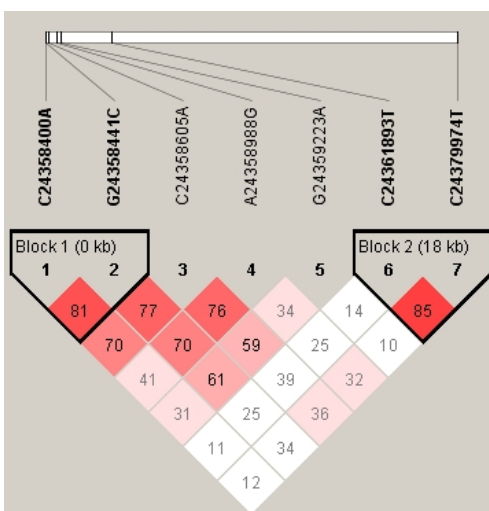

Figure 2: Linkage disequilibrium (LD) of single nucleotide polymorphisms (SNPs) at the mRNA of ATP2A3 gene in resistant and susceptible goats. Pairwise correlation (D´) values are shown between polymorphisms, which were calculated from the genotypic data of 384 goats. The haplotypes block was defined by using the default setting of the Haploview software.

Haplotype block and LD structures were generated from the seven SNPs genotyped in the ATP2A3 gene from goat populations (**Figure 2**). Pairwise correlation (D´) values are shown between polymorphisms, which were calculated from the genotypic data of 384 goats. The haplotypes block was defined by using the default setting of the Haploview software. Two pair variants, C24361893T and C24379974T from exon region as well as C24358400A and G24358441C from 3´ UTR region showed significant LD with each other with strong correlations (D´ = 85 and 81) and spanning 18 kb and 0 kb blocks, respectively.

Table 6: Frequency of haplotypes based on linkage disequilibrium (LD) in the ATP2A3 gene of Goat populations

Haplotype blocks	SNPs (N)	Haplotypes	Observations	Frequencies
Block 1 (0 Kb)	C24358400A (384) and G24358441C (384)	AG	499	0.650
		CG	202	0.263
		CC	59	0.077
		AC	8	0.010
Block 2 (18 Kb)	C24379974T (384) and C24361893T (384)	CC	409	0.533
		TT	192	0.250
		TC	150	0.195
		CT	17	0.022

N = number of individuals, SNPs = single nucleotide polymorphisms

Therefore, haplotypes were constructed from these four polymorphic sites of ATP2A3 gene (**Table 6**) and their haplotype combinations, or diplotypes. In block 1, four haplotypes were discovered in the ATP2A3 gene, where the AG haplotype showed the highest frequency (65.0%), followed by CG (26.3%), CC (7.7%) and AC (1.0%) and also four haplotypes discovered in the block 2, where the CC haplotype showed the highest frequency (53.3%),

followed by TT (25.0%), TC (19.5%) and CT (2.2%) .

Detection of relative expression of ATP2A3 gene by qRT-PCR

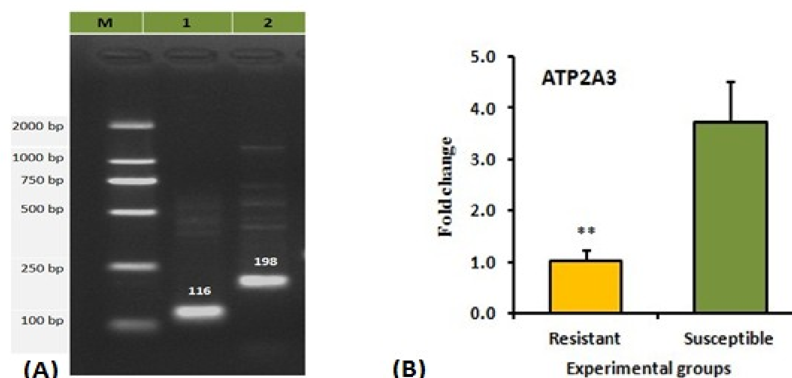

*Figure 3: The relative expression of the mRNA in ATP2A3 gene in resistant and susceptible goats. (A) The figure showing the product length at 116 and 198 bp of β-actin (housekeeping) and ATP2A3 genes, respectively. M = marker (DL 2000), 1 = β-actin housekeeping gene and 2 = ATP2A3 gene. (B) The expression value at the mRNA level was detected using qRT-PCR. The samples were normalized by β-actin housekeeping gene. The expression of the mRNA in ATP2A3 gene was highly expressed in susceptible goats compared to resistant goats. **differences were considered significant when P < 0.05 under t-test and errors bars represent standard error of the mean (SEM).*

The relative expression values in the mRNA of ATP2A3 gene in resistant and susceptible goats are shown in **Figure 3 (A) and (B)**. The relative expression of mRNA in ATP2A3 gene was higher in susceptible goats compared to that of resistant goats. The mean fold change values were significantly differed at $P < 0.05$ levels as expected similar relative fold change value from the RNA sequenced data.

Discussion

The SERCAs is a protein family possess multiple isoforms, where ATP2A3 genes encoding SERCA3 isoform (Floresperedo et al., 2016). This gene acts as a trigger in various processes of cellular activities, such as cell differentiation, muscle contraction, cell proliferation, oxidative stress, apoptosis, protein synthesis and folding (Monteith et al., 2007). Polymorphisms in the ATP2A3 gene and their associations have been identified in human gastric and colon cancer cells (Floresperedo et al., 2016). No polymorphisms in the ATP2A3 gene that may affect any disease traits in goats have been investigated. From our previous research, a number of differentially expressed genes were revealed by RNA sequences analysis using our resource YCW goat populations that are generated selected breeding program following high and low fecal eggs count (EPG) (Bhuiyan et al., 2017). Among these differentially expressed genes, ATP2A3 gene was shown higher relative log2 fold change value (4.05867) in susceptible goat compare to the resistant goat. We also validated this expression by qRT-PCR and also found 3.732 (log2) relative expression values in susceptible goat than the resistant goat in the mRNA of ATP2A3 gene (**see Figure 3**). On the basis of these results, we assumed that this gene might be one of the candidate genes for GIN infection in goat, that's why we considered it for further

polymorphism studies to identify marker-assisted selection in goat. Therefore, the candidate gene studies are one of the primary but very important techniques (Zhang et al., 2009) to determine whether a gene of interest is associated with selected disease related traits. In this study, seven unknown SNPs were identified, two of which were located in exon regions and five of which were in 3′ UTR regions of the goat ATP2A3 gene (**See Table 2**). The SNPs then genotyped in the 384 goat individuals using MALDI-TOF-MS assay. The allele frequencies showed that all those SNPs were polymorphic (MAF = 0.085 to 0.445) in these three goat breeds. As the *PIC* values classification ($PIC < 0.25$ = low polymorphism; $0.25 < PIC < 0.50$ = intermediate polymorphism; $PIC > 0.50$ = high polymorphism), the three goat populations had intermediate levels of genetic diversity (Jie et al., 2015). This result is also supported by the values of *He* and *Ne* in this study. This suggested that there was sufficient genetic diversity for selection to be effective in controlling GIN infections in these three goat populations. Genotypic and allelic frequencies, and gene diversity index in goat ATP2A3 gene indicates that the YCW goat population has lower genetic variations than that of ESB and NJY populations except for G24358988A variants, which reveals that the ESB and NJY goat population has more scope to go artificial selection than the YCW goat population (Andersson, 2001). It could be speculated that more genetic progress should be expected in ESB and NJY in terms of genetic diversity.

The SNPs C24379974T and C24361893T found in the exon regions of ATP2A3 gene were synonymous polymorphisms (**see Table 1**), which was found to have a significant association with the GIN infection traits in goat (**see Table 5**). Although synonymous mutations do not change the sequence or structure of the protein that they encode, they could affect the messenger RNA structure, splicing, stability, or folding of the protein, in where protein function will be significantly affected (Sauna, 2009). Genetic polymorphisms in the ATP2A3 gene have been positively associated with colon cancer in human (Floresperedo et al., 2016), but, no available data associated the with GIN infection traits in goat, though some studies have been carried out on other genes (Bressani et al., 2014; Alim et al., 2016; Asif et al., 2016) in goat and have revealed positive correlations with these traits. Hence, our findings of these two synonymous mutations in the goat ATP2A3 gene and its significant association with GIN infection traits recommends that these two synonymous polymorphisms could be used as a molecular marker for GIN infection traits in goat.

Likewise, the two significant mutations were detected in 3′ UTR regions (C24358400A and G24358441C) had associations with GIN infection traits (**see Tables 2 and 5**). Previous findings have shown that transcription factors can bind a few specific sites located in the intron and promoter regions of genes to influence the protein translation (Down et al., 2012) and gene expression (Myers et al., 2007; Pereyra et al., 2012). Therefore, we hypothesize that the variants identified from our study could disrupt some transcription factor-binding sites (Sheng et al., 2013) that alter ATP2A3 gene expression and affect the control of GIN infection in goat, although in-depth further studies are required.

LD analysis indicated that the C24379974T and C24361893T SNPs are a close LD pair constructed block 2 and G24358441C and C24358400A SNPs LD pair constructed block 1 (**see Figure 2**), which implies that these polymorphisms are associated with our studies GIN infection traits. Our results show that these variants of the ATP2A3 gene exert a highly additive effect on GIN infection traits. Therefore, haplotype formation from these mutations belonging to the LD blocks was consistent. Haplotype analysis is effective in LD studies to resolve unsatisfying and noisy effects than analysis of single marker which is caused by diverse marker history and statistical methods and results in monotonic and step-like breakdown of LD by recombination

(Daly et al., 2001). As our haplotype block was based on two SNPs, more tag SNPs should be typed (Zhang et al., 2009) to facilitate detecting complete haplotype blocks using different goat populations which may control GIN infection regulation. Therefore, ATP2A3 might act as a candidate gene of quantitative trait loci and significant variants in our study could be the choice of marker-assisted selection for the regulation of GIN infection in goat.

Conclusion

In this study, we described seven previously unknown polymorphisms, eight haplotypes from four variants having high LD of the goat ATP2A3 gene, as well as their significant association with GIN infection traits. These results indicate that *ATP2A3* is positively associated with GIN infection regulation, and the alleles and haplotypes of this gene may serve as genetic markers for future marker-assisted selection of GIN infection traits in goat. From qRT-PCR experiments, this gene was highly expressed in susceptible goats than the resistant goats that also indicated to a good possible candidate gene either as a major gene or as an associated with a major gene in GIN infection traits in goat. Therefore, the in depth functional study of this gene is needed to investigate the biological implication of these single nucleotide polymorphisms. To clarify further variation of the studied traits, it is recommended to do further studies on a large number of goat populations considering different genetic backgrounds.

References

Alim, M.A., Fu, Y., Wu, Z., Zhao, S.H., Cao, J., 2016. Single Nucleotide Polymorphisms of Toll-Like Receptors and Association with Haemonchus contortus Infection in Goats. Pakistan Veterinary Journal 36, 286-291.

Andersson, L., 2001. Genetic dissection of phenotypic diversity in farm animals. Nature Reviews Genetics 2, 130-138.

Asif, A.R., Ahmed, S., Du, X., Awais, M., Ijaza, N., Ansari, A.R., Yuhua, F.U., Wu, Z., 2016. Single nucleotide polymorphisms in DRB1, IGF1 and ILs associated with fecal egg count confers resistance against Haemonchus contortus infection in goats. 53.

Barrett, J.C., Fry, B., Maller, J., Daly, M.J., 2005. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21, 263-265.

Benjelloun, B., Alberto, F.J., Streeter, I., Boyer, F., Coissac, E., Stucki, S., Benbati, M., Ibnelbachyr, M., Chentouf, M., Bechchari, A., 2015. Characterizing neutral genomic diversity and selection signatures in indigenous populations of Moroccan goats (Capra hircus) using WGS data. Frontiers in Genetics 6, 107.

Bhuiyan, A.A., Li, J., Wu, Z., Pan, N., Adetula, A.A., Wang, H., Cheng, Z., Tang, X., Bhuyan, A.A., Zhao, S., 2017. Exploring the Genetic Resistance to Gastrointestinal Nematodes Infection in Goat Using RNA-Sequencing. International Journal of Molecular Sciences 18, 751.

Bickhart, D.M., Rosen, B.D., Koren, S., Sayre, B.L., Hastie, A.R., Chan, S., Lee, J., Lam, E.T., Liachko, I., Sullivan, S.T., 2017. Single-molecule sequencing and chromatin conformation capture enable de novo reference assembly of the domestic goat genome. Nature Genetics 49, 643-650.

Bishop, S.C., 2012. Possibilities to breed for resistance to nematode parasite infections in small ruminants in tropical production systems. Animal An International Journal of Animal Bioscience 6, 741-747.

Botstein, D., White, R.L., Skolnick, M., Davis, R.W., 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. American Journal of Human Genetics 32, 314.

Bressani, F.A., Tizioto, P.C., Giglioti, R., Meirelles, S.L., Coutinho, R., Benvenuti, C.L., Malagójr, W.,

Mudadu, M.A., Vieira, L.S., Zaros, L.G., 2014. Single nucleotide polymorphisms in candidate genes associated with gastrointestinal nematode infection in goats. Genetics & Molecular Research Gmr 13, 8530-8536.
Chauhan, K.K., Rout, P.K., Singh, P.K., Mandal, A., Singh, H.N., Roy, R., Singh, S.K., 2003. Susceptibility to natural gastro-intestinal nematode infection in different physiological stages in Jamunapari and Barbari goats in the semi-arid tropics. Small Ruminant Research 50, 219-223.
Daly, M.J., Rioux, J.D., Schaffner, S.F., Hudson, T.J., Lander, E.S., 2001. High-resolution haplotype structure in the human genome. Nature Genetics 29, 229.
Dodgson, J.B., Cheng, H.H., Okimoto, R., 1997. DNA marker technology: a revolution in animal genetics. Poultry Science 76, 1108.
Down, C.F., Millour, J., Lam, E.W., Watson, R.J., 2012. Binding of FOXM1 to G2/M gene promoters is dependent upon B-Myb. Biochimica Et Biophysica Acta 1819, 855.
Emara, M.G., Kim, H., 2003. Genetic markers and their application in poultry breeding. Poultry Science 82, 952-957.
Floresperedo, L., Rodríguez, G., Zarainherzberg, A., 2016. Induction of cell differentiation activates transcription of the Sarco/Endoplasmic Reticulum calcium-ATPase 3 gene (ATP2A3) in gastric and colon cancer cells. Mol Carcinog.
Frankham, Richard, 1996. Introduction to quantitative genetics (4th edn). Trends in Genetics 12, 280-280.
Howell, S.B., Burke, J.M., Miller, J.E., Terrill, T.H., Valencia, E., Williams, M.J., Williamson, L.H., Zajac, A.M., Kaplan, R.M., 2008. Prevalence of anthelmintic resistance on sheep and goat farms in the southeastern United States. J Am Vet Med Assoc 233, 1913-1919.
Jie, W., Li, G.W., Elzo, M.A., Yan, L.J., Chen, S.Y., Jia, X.B., Lai, S.J., 2015. A novel single nucleotide polymorphism of the POU1F1 gene associated with meat quality traits in rabbits. Annals of Animal Science 15, 611-620.
Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. Oxford University Press.
Livak, K.J., 2008. Analyzing real-time PCR data by the comparative C(T) method. Nature Protocols 3, 1101.
Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25, 402-408.
Malviya, H.C., Patnaik, B., Tiwari, H.C., Sharma, B.K., 1979. Measurement of the blood loss caused by Haemonchus contortus infection in sheep. Indian Veterinary Journal.
Mckenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Research 20, 1297-1303.
Monteith, G.R., Mcandrew, D., Faddy, H.M., Roberts-Thomson, S.J., 2007. Calcium and cancer: targeting Ca2+ transport. Nature Reviews Cancer 7, 519-530.
Myers, R.L., Airey, D.C., Manier, D.H., Shelton, R.C., Sanders-Bush, E., 2007. Polymorphisms in the regulatory region of the human serotonin 5-HT2A receptor gene (HTR2A) influence gene expression. Biological Psychiatry 61, 167-173.
Nei, M., Roychoudhury, A.K., 1974. SAMPLING VARIANCES OF HETEROZYGOSITY AND GENETIC DISTANCE. Genetics 76, 379-390.
Pereyra, S., Velazquez, T., Bertoni, B., Sapiro, Rossana, 2012. Rapid multiplex high resolution melting method to analyze inflammatory related SNPs in preterm birth. Bmc Research Notes 5, 69-69.
Prasad, A., Nasir, A., Rakesh, R., Saini, V.K., 2016. Identification of immunoreactive polypeptides in'Haemonchus contortus' by immunoprecipitation during prepatent infection in sheep. Journal of Animal Research 6, 61.
Roepstorff, A., Nansen, P., 1998. Epidemiology, diagnosis and control of helminth parasites of swine. Fao

Animal Health Manual.
Sauna, Z.E., 2009. Silent (Synonymous) SNPs: Should We Care About Them? Methods in Molecular Biology 578, 23.
Sheng, Q., Cao, D., Zhou, Y., Lei, Q., Han, H., Li, F., Lu, Y., Wang, C., 2013. Detection of SNPs in the Cathepsin D Gene and Their Association with Yolk Traits in Chickens. PLoS One 8, e56656.
Sun, Y., Liu, R., Lu, X., Hu, Y., Zhao, G., Zheng, M., Chen, J., Wang, H., Wen, J., 2013. Associations of polymorphisms in four candidate genes with carcass and/or meat-quality traits in two meat-type chicken lines. Animal Biotechnology 24, 53-65.
Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2014. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S.. MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Mol Biol Evol 28: 2731-2739. Molecular Biology & Evolution 28, 2731-2739.
Vignal, A., Milan, D., Sancristobal, M., Eggen, A., 2002. A review on SNP and other types of molecular markers and their use in animal genetics. Genetics Selection Evolution,34,3(2002-05-15) 34, 275-305.
Zhang, Z.R., Liu, Y.P., Yao, Y.G., Jiang, X.S., Du, H.R., Zhu, Q., 2009. Identification and association of the single nucleotide polymorphisms in calpain3 (CAPN3) gene with carcass traits in chickens. Bmc Genetics 10, 10.