DNA Sequence Variation within STAT1 Gene of Gaolao Cattle

P. M. Khandare¹, D. S. Kale¹*, P. G. Koringa², D. V. Patil¹ and A. P. Dhok³

¹Department of Animal Genetics and Breeding, Nagpur Veterinary College, Maharashtra Animal and Fishery Sciences University (MAFSU), Nagpur, India
²Department of Animal Biotechnology, College of Veterinary Science and A.H., Anand Agricultural University, Anand, India
³Department of Animal Nutrition, Nagpur Veterinary College, Maharashtra Animal and Fishery Sciences University (MAFSU), Nagpur, India

*Corresponding author

Abstract

The current research was designed to study DNA sequence variation within Signal transducer and activator of transcription 1 (STAT1) gene and to find association of STAT1 gene variants (if any) with milk traits in Gaolao cattle. DNA was isolated from around 40 Gaolao cows from farmers herd and farm in breeding tract. The STAT1 gene regions of 314 bp, 850 bp and 154 bp fragments were amplified using STAT1G1, STAT1G2 and STAT1G4 primers respectively. PCR-RFLP analysis of 314 bp 3' UTR fragment of STAT1 gene using PglI restriction enzyme, revealed the CC genotype in Gaolao cattle indicating monomorphic status of the locus. The alignment analysis of 222 bp STAT1G1 sequence revealed SNP T-C at 221st position as polymorphic site. STAT1G2 primer amplified 783 bp sequence which revealed 03 polymorphic sites as SNP G-T at 724th position, SNP A-G at 730th position and SNP C-G at 736th position. The polymorphic sites were identified in STAT1 gene sequence indicating sequence variation in the STAT1 gene regions in Gaolao Cattle. The identified polymorphic sites may be validated with bigger phenotypic dataset and the results of which may prove significant for identification of markers related with milk production traits.

Keywords
Gaolao Cattle, PCR-RFLP, STAT1, SNP, Sequencing

Article Info
Accepted: 22 February 2020
Available Online: 10 March 2020

Introduction

India is bestowed with indigenous cattle breeds resistant to adverse climatic conditions, disease and thermal stress with fair milk yielding ability. In line with this Gaolao cattle are fair milk yielder well adapted to adverse climatic conditions and are culturally important cattle breed of Vidarabha region of Maharashtra. Genetic adaptation is characteristic with heritable basis which may aid to improve productivity, disease resistance, reproductive and survival traits of cattle caused by domestication and natural...
evolution (Vander et al., 2013). Current research in dairy genomics, are revealing genes and targets for genomic selection (Clancey et al., 2019) and its use in increasing genetic gain for milk production traits. STAT1 candidate gene is important because of its involvement in the development and differentiation of the mammary glands (Watson et al., 2008).

The location of STAT1 transcription factor gene is on the chromosome 2 between interval 60 to 63 cM (Band et al., 2000). It has been proved that, the STAT candidate gene influences milk composition, embryonic survival, conception rate, fertilization rate and calving interval of dairy animals (Khatib et al., 2009; Klover et al., 2010).

Various researchers have reported the association of SNP markers within STAT1 with milk production traits in dairy cattle (Mosig et al., 2001; Li et al., 2014; Cobanoglu et al., 2016) and concluded that STAT1 gene might be used as a potential candidate gene to improve milk yield and milk fat and protein contents in dairy cows breeding programs.

Cobanoglu et al., (2006) reported allele C of STAT1 gene increases fat and protein percentage of milk, and claimed that genotypes CC and CT were associated with higher milk, fat and protein yield in Holstein dairy cattle.

However, such types of studies in Gaolao cattle were relatively scanty. STAT1 candidate gene polymorphism and its strong relevance with milk production traits can be exploited for its use in marker-assisted selection in breeding. Keeping the above facts in view, the present study was undertaken to study variation in STAT1 gene with the objectives to study the DNA polymorphisms with STAT1 gene in Gaolao cattle genome using PCR-RFLP and DNA sequencing tools & to study the association of STAT1 gene variants (if any) with milk production traits in Gaolao cattle.

### Materials and Methods

#### Experimental animals and data

The purebred Gaolao cattle were surveyed in the breeding tract especially in villages of Ashti, Arvi and Karanja tehsils of Wardha district and blood and test day milk was collected from minimum of 40 Gaolao cows. A quantity of 5 ml of blood was collected from experimental Gaolao cows using a sterile needle under aseptic conditions. Around 20 ml milk samples were collected from individual cows which were processed for estimation of milk components like Fat%, protein%, SNF% and lactose%.

All the experimental protocol were duly approved by the Institutional Animal Ethics Committee.

#### DNA extraction & quality check

The phenol, chlororform-isooamylalcohol extraction protocol was used for extraction of DNA (Sambrook and Russel, 2001) with some modifications. The Nano drop reading showing ratio of 260A/280A wavelengths between 1.7 to 2.0 and DNA concentration above 50ng/µl were selected for use in PCR.

The quality and the purity of the genomic DNA were checked on 0.8% agarose gel electrophoresis based on its dense and compact appearance under GelDoc system (Bio-Era). All the genomic DNA samples were stored as stock solution at -20°C in freezer (Celfrost). A set of the working solution of 50 ng/µl concentrations for each genomic DNA was prepared for its use as template DNA in PCR reaction.
PCR primers

The reported sequences of primers (Cobanoglu et al., 2006) were used for amplification of 3’ UTR of STAT1 gene using primer STAT1G1. The primers were designed for particular exon region for STAT1G2, STAT1G3 and for UTR with respect to STAT1G4 primer of STAT1 gene (Table 1). Primers for STAT1 gene were designed based on the available sequence for Bos taurus from NCBI GenBank (Accession No.: NM_001077900) using Primer3 software (http://bioinfo.ut.ee/primer).

Primer dissolution and dilution

Oligos were supplied in freeze dried powder form were reconstituted in nuclease free water to the volume (in µl) equivalent to the mass (µg) of primer to create 100 µM/µl of stock solution and further diluted to give the final concentration. The primer STAT1G1 of 50 ng/µl concentration was prepared; however for the other primers STAT1G2, STAT1G3 and STAT1G4, the concentration was 10 pM/µl for its use in PCR reaction.

Polymerase chain reaction

PCR conditions were optimized for each primer with regard to annealing temperatures with reference to Tm of forward and reverse primers and also the concentration of primers, dNTPs, template DNA, etc. The volume of 25µl PCR reaction mixture for STAT1G1, STAT1G2 and STAT1G4 primers with different quantity and concentration were optimised (Table 3). A master mix was prepared excluding template DNA. Thermal cycling conditions like annealing temperature was optimized as 53ºC, 57ºC & 54.8ºC for the primers STAT1G1, STAT1G2 and STAT1G4 respectively for precise amplification of STAT1 gene regions. Gene fragment of 314 bp size (Figure 1) of STAT1 gene were amplified by using reported primer STAT1G1 (50 ng/µl) and designed primer of STAT1G2 and STAT1G4 (10pM) of 850 bp (Figure 3) and 154 bp product size (Figure 4) respectively were amplified. The products were checked by electrophoresis using 1.5% of agarose gel and the combinations giving the best amplification were used for further studies.

RFLP analysis using restriction enzyme PagI & STAT1 PCR product

PagI (NEB) is one of the restriction enzymes isolated from the E.coli strain from Bacillus species. The PagI enzyme cleaves the DNA at the positions where the TCATGA sequence which is recognition sequence. The PCR products were digested with PagI restriction endonuclease at 37ºC for 4 hours with the composition of digestive mixture (Table 2). The restricted digested PCR product were loaded for gel electrophoresis on 2.5 to 3% agarose gel along with 50 bp DNA ladder and electrophoresed for 30 to 120 minutes at 50 volts using 1X TBE buffer.

Direct DNA sequencing & sequence analysis using bioinformatics tools

The PCR products obtained using STAT1G2 and STAT1G4 primers were randomly selected for direct sequencing to Eurofins Genomics India Pvt. Ltd., Bangalore. The obtained DNA sequences were edited and analyzed using various bioinformatics tools (BioEdit Software & online tools viz; BLAST, ClustalW etc) for detection of nucleotide substitutions (if any).

Results and Discussion

In current investigation, the STAT1 candidate gene which was less widely reported but significant candidate gene exhibiting association with economic traits like milk
production and reproduction in dairy animals was selected for DNA sequence variation study. The STAT1 gene was studied for DNA sequence variation using PCR-RFLP and DNA sequencing tool. The findings of the current study are presented and discussed in stepwise manner under the following heads and subheads.

**STAT1G1-PagI PCR-RFLP of STAT1 gene in Gaolao cattle**

The PCR product of 314 bp fragment of STAT1 gene representing 3' UTR was digested with PagI restriction enzyme. The restriction fragments were separated in 3% agarose gel and visualized in GelDoc System for digestion results. However, all the experimental Gaolao cattle exhibited single restriction pattern designated as CC genotype indicating monomorphic status of this locus.

Restriction digestion results documented under GelDoc revealed two fragments of size 201 bp and 113 bp designated as CC genotype (Figure 2). However, Cobanoglu *et al.*, (2006) studied the STAT1-PagI locus in 639 Holstein dairy cattle and reported the frequency of CC genotype as 0.47 and CT genotype as 0.53 respectively. They also reported the frequency of C allele as 0.66 in Cooperative Dairy DNA Repository and University Farm resource populations and found that allele C was associated with significant increases in milk fat and protein percentages. Hengameh and Mojtaba (2016) investigated SNPs within 314 bp STAT1 gene in Brown Swiss dairy cattle at STAT1-PagI locus which revealed genotype frequencies of TT, TC and CC as 0.23, 0.25 and 0.52 respectively.

They also reported the frequency of C allele as 0.66 in Cooperative Dairy DNA Repository and University Farm resource populations and found that allele C was associated with significant increases in milk fat and protein percentages. Hengameh and Mojtaba (2016) investigated SNPs within 314 bp STAT1 gene in Brown Swiss dairy cattle at STAT1-PagI locus which revealed genotype frequencies of TT, TC and CC as 0.23, 0.25 and 0.52 respectively.

The T and C allele frequency were 0.355 and 0.645 respectively. The results of present study show that the T allele is minor frequency allele and C allele has positive effect on milk, fat and protein yield traits. Later, Cobanoglu *et al.*, (2016) carried out genotyping at STAT1-PagI locus in 472 Holstein cattle and found the frequency of CC,CT & TT genotype as 0.49,0.40 and 0.11 respectively. In 283 Jersey cattle the frequency CC,CT and TT genotype was reported as 0.44,0.52,0.04 respectively. They found that CT genotypes were found significant for corrected 305 days milk yield. However, TT genotypes were 2.07 kg higher for test-day milk yield, 0.13 kg for fat yield and 0.07 kg for protein yield.

Ardicli *et al.*, (2018) studied STAT1-PagI gene polymorphisms in 168 purebred Holstein-Friesian cows and reported genotype frequencies of CC, CT & TT as 0.45,0.36 and 0.19 respectively. In total of 317 Holstein dairy cows, Askari *et al.*, (2013) reported seven genotypes as DD (89), BB (25), CC (11), AC (15), BC (9), CD (164) and BD (4) at STAT1-PagI locus. They reported frequency of A, B, C, and D alleles as 0.021, 0.101, 0.332 and 0.546, respectively. The results of association study indicated STAT1-PagI locus was highly related (P < 0.05) with fat%; and found that DD was superior genotype as compared to other groups. Due to these results they concluded STAT1 marker should be considered for milk fat percentage selection in Holstein dairy cattle. Even in the buffaloes the fragment of 314 bp was screened for polymorphism using other restriction enzyme, the result was monomorphic. Kumar *et al.*, (2015) explored 3' UTR region of same 314 bp fragment of STAT1 gene in Murrah, Gojri and Chattisgarhi buffaloes using HaeIII restriction enzyme , which was resulted in to monomorphic locus.

The obtained edited STAT1G1 amplified sequence of 222 bp size of Gaolao cattle was aligned with reference sequence (Accession
No. XM_025000697.1) using CLUSTAL Omega tool which revealed polymorphic sites as SNP T-C at 221st position. DNA sequencing studies carried by other researchers in bovine STAT1 gene reported SNPs. Similarly, Cobanoglue et al., (2006) used pooled genomic DNA sequencing & identified an SNP (C/T) at position 213 in the EST corresponding to STAT1 Gene in Holstein dairy Cattle. In order to know the sequence homology of STAT1G1 amplified 222 bp 3’UTR region of Gaolao gene with other species BLAST analysis was performed & developed BLAST tree view which revealed sequence identity of 99.55 % with Bos indicus, Bos taurus, Bubalus bubalis, 99.10% with Bison bison, 93.81% & 92.92% with Ovies aries and Capra hircus respectively and 80.95% with Camel.

The tree was produced using BLAST pairwise alignments (https://blast.ncbi.nlm.) between query sequence of 222 bp obtained using STAT1G1 primers in Gaolao cattle with database sequences of other related species.

BLAST is useful tool to know the sequence variation hence, Cobanoglue et al., (2006) used BLAST which revealed 100% similarity between this EST and the predicted sequence of the 3’untranslated region (UTR) of the STAT1 gene (GenBank accession number XM_872927). Thus, SNP C/T corresponded to position 3141 in the 3’UTR of the predicted sequence of STAT1.

DNA sequence analysis of STAT1G2 amplified (850 bp) 783 bp STAT1 amplicon sequence in gaolao cattle

The edited STAT1G2 amplified sequence of 783 bp size of Gaolao cattle and Bos taurus sequence alignment was performed (CHR2:79523839-79525254) using CLUSTAL tool which resulted in to three variable sites as SNP G-T at 724th position, SNP A-G at 730th position and SNP C-G at 736th position (Figure 5). In order to know the sequence homology of STAT1G2 amplified 783 bp region of Gaolao cattle gene with other species BLAST analysis was performed & developed BLAST tree view which revealed sequence identity of 98.36 % with Bos indicus, Bos taurus and Bison bison, 97.54% with Bubalus bubalis,Capra hircus and Bos mutus, 96.72% with Ovies aries and with 93.44% Sus scrofa respectively. Dendrogram was generated using BLAST alignments tool for query sequence of 783 bp obtained using STAT1G2 primers in Gaolao cattle with DNA database sequences of species (Figure 6).

Similarly intronic and exonic regions of STAT1 gene were studied by other researchers for identification of SNPs. Deng et al., (2016) studied STAT1 gene to evaluate SNPs and their significant associations with milk production traits in 192 buffaloes. They identified 18 SNP in the buffalo STAT1 gene, including 15 intronic mutations and 3 exon mutations & revealed that SNP13 located in exon 10 were significantly associated with the milk production traits.

The study has provided evidence that buffalo STAT1 gene due to its association with milk production traits and can be used as a marker for selection & breeding decisions. The obtained DNA sequence (783 bp) using STAT1G2 primers in Gaolao cattle was analyzed using Translate (https://web.expasy.org/translate/) tool to predict the protein sequence which resulted three frames out of that the frame exhibiting longest ORF was selected which resulted in to 80AA. The obtained protein sequence was aligned using BLASTP (version 2.2.31+) programme with UniProtKB database which revealed the identity of the query sequence with uncharacterized protein A0A1A9YBG1.
Table 1 Name, Primer Sequence, product size and gene region used to study STAT1 gene sequence variation in Gaolao Cattle

<table>
<thead>
<tr>
<th>Name of Primer/ region</th>
<th>Primer sequence</th>
<th>Product Size (bp)</th>
<th>RE/Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>STAT1G1/ 3’ UTR</td>
<td>FP: 5’-GCCTCAAGTTTGCCAGTGCC-3’ RP: 5’-GGCTCCCTTGATAGAAGACTGT-3’</td>
<td>314 bp</td>
<td>PagI (Cobanoglu et al., 2006)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STAT1G2/ Exon</td>
<td>FP : 5’-GCAGTTGAACCTACACAAAA-3’ RP : 5’-TTTAGGGGCMTCATAGTTCCA-3’</td>
<td>850 bp</td>
<td>Designed</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STAT1G3/ Exon</td>
<td>FP: 5’-TGATGAACGTGGAGTGGA-3’ RP: 5’-GAATGAGCTGCAAGAAGAGG-3’</td>
<td>242 bp</td>
<td>Designed</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STAT1G4/ 3’ UTR</td>
<td>FP: 5’-CGTGGGCTTGAGCTGAATT-3’ RP: 5’-GCCAGCATCAATAAGCCATA-3’</td>
<td>154 bp</td>
<td>Designed</td>
</tr>
</tbody>
</table>

Table 2 Optimized Restriction Enzyme digestion reaction mixture for digestion of 314 bp STAT1G1 amplified PCR product in Gaolao Cattle

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Reaction components</th>
<th>Quantity</th>
<th>Final Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>1X cut smart buffer</td>
<td>2.5 µl</td>
<td>1X</td>
</tr>
<tr>
<td>2.</td>
<td>RE (500U/µl)</td>
<td>0.2 µl</td>
<td>0.2 U</td>
</tr>
<tr>
<td>3.</td>
<td>PCR product</td>
<td>10 µl</td>
<td>50ng/µl</td>
</tr>
<tr>
<td>4.</td>
<td>MBG water</td>
<td>7.3 µl</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>20 µl</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 3 PCR reaction volume containing optimized different concentration and quantity for STAT1G1 and STAT1G2 of STAT1 gene in Gaolao cattle

<table>
<thead>
<tr>
<th>Sr No.</th>
<th>Name of the PCR Components</th>
<th>Quantity (µl) (STAT1G1)</th>
<th>Quantity (µl) (STAT1G2)</th>
<th>Final Concentration</th>
<th>Quantity (µl) (STAT1G4)</th>
<th>Final Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>PCR Buffer with MgCl$_2$ (10x)</td>
<td>2.5</td>
<td>2.5</td>
<td>10X</td>
<td>12.5</td>
<td>ProMega Master Mix</td>
</tr>
<tr>
<td>2.</td>
<td>dNTPs</td>
<td>1.0</td>
<td>1.5</td>
<td>200µM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>Forward Primer</td>
<td>0.5</td>
<td>1.0</td>
<td>10pM</td>
<td>1.0</td>
<td>10pM</td>
</tr>
<tr>
<td>4.</td>
<td>Reverse Primer</td>
<td>0.5</td>
<td>1.0</td>
<td>10pM</td>
<td>1.0</td>
<td>10pM</td>
</tr>
<tr>
<td>5.</td>
<td>Taq DNA Polymerase (5 U/µl)</td>
<td>0.3</td>
<td>0.3</td>
<td>1U/µl</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6.</td>
<td>MBG Water</td>
<td>18.7</td>
<td>16.7</td>
<td>-</td>
<td>8.5</td>
<td>-</td>
</tr>
<tr>
<td>7.</td>
<td>Template DNA (50ng/µl)</td>
<td>1.5</td>
<td>2.0</td>
<td>50ng/µl</td>
<td>2.0</td>
<td>50ng/µl</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>25µl</td>
<td>25µl</td>
<td>-</td>
<td>25µl</td>
<td>-</td>
</tr>
</tbody>
</table>
**Figure 1** STAT1 gene PCR product (size – 314 bp) amplified using STAT1G1 Primer in Gaolao cattle and resolved in 1.5% Agarose gel electrophoresis.
Where, L-50 bp DNA Ladder G1 to G7 – Gaolao Cattle Numbers

**Figure 2** STAT1 gene polymorphism (STAT1-Pag1) in Gaolao cattle using Pag/restriction Enzyme for digestion of 314 bp PCR product amplified using STAT1G1 primer and visualized in 3% agarose gel electrophoresis. Where, LD = Ladder 50 bp; GS 3-16 = Gaolao cattle number; CC = 201 bp genotype
Figure 3 STAT1 gene PCR Product (Size - 850 bp) amplified using STAT1G2 primer in Gaolao cattle and resolved in 1.5% Agarose gel electrophoresis. Where, L= 50bp ladder; G1-G8 = Gaolao cattle Numbers

Figure 4 STAT1 gene PCR product (size – 154 bp) amplified using STAT1G4 primer in Gaolao cattle and resolved in 1.5% Agarose gel electrophoresis. Where, L = 50 bp Ladder; G1-G8 = Gaolao Cattle Numbers
However, the obtained edited STAT1G4 amplified sequence of 87 bp size of Gaolao cattle was aligned with reference sequence (Accession No. XM_025000697.1) using CLUSTAL O alignment tool which revealed absence of polymorphic sites. In conclusion, it is to state that, various researchers has reported SNP markers within STAT1 gene.
and indicated status of STAT1 gene as a significant candidate gene for its use in selection decisions for genetic improvement of cattle for productivity. The 3' UTR region of STAT1 gene of 314 bp size was analysed using PCR-RFLP tool with pagI restriction enzyme. Restriction digestion analysis at STAT1G1- pagI locus exhibited CC genotype in all the experimental animals indicating fixation of C allele & monomorphic status of this locus in the population. The STAT1 gene in Gaolao cattle exhibited 04 SNPs as T-C at 221st position, SNP G-T at 724th position, SNP A-G at 730th position and SNP C-G at 736th position indicating presence of DNA sequence variation.

As per the results of previous studies allele C has been associated to increases fat and protein percentage of milk in Holstein cattle. However, the current study needs to conduct in large number of animals with accurate milk trait data to reveal DNA sequence variants related with milk production traits for their use in selection and genetic improvement.

**Acknowledgement**

The author acknowledges the financial support by the funding agency SERB-DST, GOI, New Delhi under Core Research Grant Scheme (File No. EMR/2017/000323 dt.29/08/2018) under Project Investigator, Dr.D.S.Kale, Assistant Professor at Deptt. of Animal Genetics and Breeding, Nagpur Veterinary College, MAFSU, Nagpur. The help rendered by members of Gaolao Breeder’s Association in Survey and phenotyping of Gaolao cattle in the breeding tract is duey acknowledged.

**References**

Ardicli, S., B. Soyudal, H. Samli, D. Dincel and Balci, F. 2018. Effect of STAT1, OLR1, CSN1S1, CSN1S2, and DGAT1 genes on milk yield and composition traits of Holstein breed. Revista Brasileira de Zootecnia. 47: 247.


Hengameh, V. and Mojtaba, H. M. 2016. Association of single nucleotide


How to cite this article: