Molecular characterization and phylogeny based analysis of intron I sequence of Myostatin (Mstn) gene in Iranian Makuei sheep breed

Mohammad Farhadian1*, Ali Hashemi2

1Young Researchers and Elite Club, Khoy Branch, Islamic Azad University, Khoy, Iran
2Department of Animal Science, Faculty of Agriculture, Urmia University, Urmia, Iran
*Corresponding author: mohammad.farhadian@tabrizu.ac.ir

Abstract
Myostatin (MSTN), a negative regulator of skeletal muscle development, acts as a potential candidate gene used to increase muscle mass. Likewise, sheep MSTN gene has an important role in meat production. MSTN is made up of 376 amino acids, and is synthesized as a precursor protein. To investigate the MSTN in Iranian native Makuei sheep, a polymerase chain reaction-single strand conformational polymorphism (PCR-SSCP) analysis was used. Genomic DNA was isolated from the blood samples. A 417-bp of MSTN intron I segment was amplified using locus-specific primers. Four SSCP patterns were identified and nucleotide sequencing of the Makuei sheep MSTN gene was done and registered in the NCBI GenBank with “kJ526625” number. Three novel single nucleotide polymorphisms (SNPs) were found in the samples. These SNPs are found in 224bp, 226bp and 242bp locations. Accordingly three substitutions (c.224C>T; c.226A>G; c.242G>T) were observed in the intron 1 region of Mstn gene. The effects of the observed SNPs on breeding values of some biometric traits were investigated and the substitution of c.226A>G was found to be associated with heart girth (HG) and leg circumference (LC). Phylogenetic analysis, based on the nucleotide sequences indicated similar evaluation with the GenBank reference sequences. It seems that the observed polymorphisms of the ovine MSTN gene are associated with HG and LC traits.

Key words: myostatin (MSTN), Makuei sheep, phylogeny, polymorphism

Myostatin (MSTN) gene in ovine is located on chromosome 2 and consists of three exons and two intron regions (McPherron and Lee, 1997), similar to some animals such as cattle (Grobe et al., 1997), pig (AY208121), buffalo (AH013313), zebra fish (AY323521), gilthead sea bream (AF258447), chicken (AF346599) and house mouse (AY204900) (Grobe et al., 1997).

Mutation in this gene causes muscular hypertrophy ending up to developing muscular mass (McPherron and Lee, 1997). Presence of the muscular hypertrophy
alleles leads to muscular hypertrophy phenotype. However, the mutation does not have to be in the homozygous condition in order for the effect to be detected.

The animals which carry a single copy of the muscular hypertrophy mutation allele from crossbred Belgian Blue, or crossbred Piedmontese sire, had increased longissimus muscle area and retail yield, and reduced external and intramuscular fat deposition compared with animals receiving no copies thereof (Casas et al., 1998).

MSTN gene is highly conserved among all species. Its variation in some mammalian species such as mice (McPherron and Lee, 1997), cattle (Grobet et al., 1997; Dunner et al., 2003), humans (Schuelke et al., 2004), dogs (Mosher et al., 2007), pigs (Stinckens et al., 2008) and sheep (Kijas et al., 2007; Boman and Vage, 2009; Johnson et al., 2009; Hickford et al., 2010; Han et al., 2013) has been reported in association with muscular mass.

Polymerase chain reaction-singlestrand conformational polymorphism (PCR-SSCP) analysis of intron-1 is used to identify genetic variation of the MSTN gene. In an experiment on 110 Lory sheep in Lorestan province, for example, five unique SSCP patterns (genotypes) and five alleles were located. No statistically significant difference was observed for BW (body weight) trait among the genotypes of the population. However, there was a statistically significant difference in WW (weaning weight) and BW6 traits among the genotypes of the population (P<0.05). The chi-square test showed that the population has deviated from Hardy-Weinberg equilibrium for the studied locus (P<0.05) (Sepahvand et al., 2014).

Analysis of the intron I region of this gene, using single-strand conformational polymorphism (SSCP), revealed 5 allelic variants in the New Zealand Romney sheep breed (Hickford et al., 2010). In another study on MSTN gene polymorphism in exon 3 (using PCR-RFLP), intron 1 and intron 2 (using PCR-SSCP) and their association with yearling weight records in Zel sheep were done. Under RFLP method, all samples showed mm genotype. Under SSCP method, intron 1 was also monomorphic, but intron 2 was polymorphic (Dehnavi et al., 2012).

Additional SNPs associated with muscling have also been identified in the promoter and intron 2 regions (Kijas et al., 2007). In total, 23 ovine MSTN SNPs have been identified (Clop et al., 2006; Kijas et al., 2007; Zhou et al., 2008). To date, three allelic variants have been reported in a 473-bp section of the exon 1- intron 1 region (Zhou et al., 2008) and four haplotypic variants have been constructed through the coding sequence using SNPs positioned at 41, +4036 and +6223 (Kijas et al., 2007).

The main purpose of sheep husbandry has been to produce meat which is needed according to the increasing worldwide demands. Considering pastures being exhausted due to overuse, it seems both wise and necessary to produce more meat per sheep. More opportunities lie within the sheep itself, beneath muscle-related mutations, which alongside breeding value of biometric traits can serve us well if identified and utilized purposely. Therefore, the objective of the present study has been to detect potential association of variation in the Makuei sheep MSTN gene with breeding value of biometric traits and also sequence analysis of intron 1 region and to compare it with other sheep based on nucleotide sequence.
Material and methods

Breed description: sampling and DNA extraction
Two hundred Makuei sheep were randomly selected, and their biometric traits (HW: height at withers; HR: height at rump; BL: body length; HG: heart girth; LC: leg circumference) records were obtained from the Makuei sheep breeding center. Blood samples from all selected sheep were collected. The whole blood was preserved in ethylenediamine tetra acetic acid (EDTA)-coated tubes and stored at −20°C. Genomic DNA was extracted using a genomic DNA purification kit (Fermentas, USA) according to manufacturer instructions.

PCR amplifications and detection of genetic variations using SSCP
The intron 1 region of the MSTN gene was amplified using two primers (5′ GAAACGGGTATTACCATGC-3′) and (5′-CATATTTCAGGACCAATG-3′) targeting a fragment of 417 bp designed at NCBI website (http://ncbi.nlm.nih.gov/). The PCRs were run in a final volume of 50 µL using a PCR Master Mix kit (CinnaGen Inc., Tehran, Iran) containing 2.5 U Taq DNA polymerase in reaction buffer, 4 mM MgCl2, 50 µM of each dNTP, 0.5 µM of each primer, and 100 ng extracted DNA as a template. The thermal cycling profile consisted of an initial denaturation step of 94°C for 2 min, 35 cycles of denaturation at 94°C for 45 seconds; annealing at 52°C for 45 seconds and extension at 72°C for 45 seconds followed by a final extension at 72°C for 2 min. Products of amplification were recognized by electrophoresis on 2% agarose gel stained with ethidium bromide. PCR products were mixed with 8 µL denaturing loading dye [95% (w/v) deionized formamide, 0.05% (w/v) xylene cyanol, 0.05% (w/v) bromophenol blue, and 0.02 M EDTA] in a total volume of 15 µL. The mixture was denatured at 95°C for 5 min and then snap-chilled on ice (Pipalia et al., 2004). The total volume was electrophoresed on 8% polyacrylamide gel, as described by Herring et al. (1982). The electrophoresis was run in 0.5X TBE buffer at room temperature (18°C) with a constant 200 V for 3 h. Polyacrylamide gels were stained using silver nitrate according to the protocol described by Herring et al. (1982).

Estimation of breeding values for biometric traits and statistical analysis
A total of 10515 body measurement records from 3062 individual progenies of 489 sires and 3013 dams over 14 years (2000 to 2013) were used in the present study. Animals’ biometric traits were measured at yearling (12-month-old) age.

Data on biometric traits (HW: height at withers; HR: height at rump; BL: body length; HG: heart girth; LC: leg circumference) were obtained from the national sheep recording system. The following linear model was used to calculate breeding values using WOMBAT software (Meyer, 2007):

\[ Y_{ijklmn} = \mu + YR_i + SX_j + BT_k + AD_l + AN_m + e_{ijklmn} \]

where:
\[ YR_i \] – the fixed effect of \( i \)th year (1–14),
$S_{X_j}$ – the fixed effect of $j^{th}$ sex (1–2),
$B_{T_k}$ – the fixed effect of $k^{th}$ birth type (1–3),
$A_{D_l}$ – the fixed effect of $l^{th}$ month of birth (1–7),
$A_{N_m}$ – the fixed effect of $m^{th}$ level of the random additive genetic effect ($m =$ number of animal for each trait) and $e_{ij}$ is residual effects.

The association of genotypes with breeding values (BV) of biometric traits was investigated using the GLM procedure of SAS 9.1 software. The following linear equation was used to perform the association analysis:

$$BV_{ij} = \mu + G_i + e_{ij}$$

where:
$BV_{ij}$ – is predicted breeding value of biometric trait for $j^{th}$ individual with $i^{th}$ genotype,
$\mu$ – overall mean for each trait,
$G_i$ – is the effect of the $i^{th}$ SSCP genotype,
$e_{ij}$ – is the random error effect.

The Duncan method was used for comparison of the means.

**Sequence information**

PCR products from four different patterns of the MSTN gene were subjected to DNA sequencing. Each PCR product (10 µL) and 15 µL with 5 µM MSTN up primer were sent to CinnaGen for PCR cleanup and sequencing. MEGA 6.0 and BioEdit software were used to obtain information on the various patterns sequenced. A multiple-sequence alignment was performed using BioEdit software (http://www.mbio.ncsu.edu/bioedit/bioedit.html/). Sequence conservation along alignments was visualized as an entropy plot, which is a measure of lack of predictability for an alignment position. A score of zero stands for complete conservation (completely predictable), whereas higher positive values indicate greater diversity between the sequences. A phylogenetic tree based on intron I region was generated using Neighbor-Joining method from MEGA version 6.0 (Tamura et al., 2007).

**Results**

**PCR amplification of intron I of MSTN gene and SSCP analysis**

The intron I of the MSTN gene with 417bp in length was successfully amplified by the specific primer. The allelic variation in the MSTN gene was examined using a PCR-SSCP method. Four SSCP genotypes were identified in the intron I of Makuei sheep of MSTN gene (Figure 1). The frequencies of the observed genotypes were 0.413, 0.293, 0.130, and 0.163 for TAT, CGT, TGG, and TGT, respectively.
Molecular characterization of MSTN gene in Makuei Sheep

Figure 1. SSCP polymorphism of Makuei sheep MSTN gene

**Associations with biometric traits**

Table 1 presents data related to association of the observed genotypes of intron I of MSTN gene with breeding value of biometric traits such as height at withers (HW), height at rump (HR), body length (BL), heart girth (HG) and leg circumference (LC). The general linear mixed model revealed that the TAT genotype was associated with HG and LC traits, but no association was found between other observed genotypes (CGT, TGG and TGT) with HG and LC traits. Plus, the effects of TAT, CGT, TGG and TGT genotypes on HW, HR and BL traits were non-significant (Table 1).

Table 1. Effect of the MSTN gene genotype on BV of biometric traits in Iranian Makuei sheep breed

<table>
<thead>
<tr>
<th>Trait</th>
<th>GENOTYPES</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TAT</td>
<td>CGT</td>
</tr>
<tr>
<td>HW</td>
<td>0.011±0.001</td>
<td>0.016±0.001</td>
</tr>
<tr>
<td>HR</td>
<td>0.782±0.068</td>
<td>0.876±0.084</td>
</tr>
<tr>
<td>BL</td>
<td>0.051±0.003</td>
<td>0.049±0.002</td>
</tr>
<tr>
<td>HG</td>
<td>0.381 a±0.046</td>
<td>0.191 bc±0.051</td>
</tr>
<tr>
<td>LC</td>
<td>0.435 a±0.039</td>
<td>0.214 b±0.062</td>
</tr>
</tbody>
</table>

HW: height at withers; HR: height at rump; BL: body length; HG: heart girth; LC: leg circumference.

**Analysis of Intron I of MSTN gene sequence**

Sequence analysis was done using the MEGA software. A typical sequence representing the four different observed genotypes consisted in average of 30.5% thiamine (T), 16.3% cytosine (C), 31.4% adenine (A) and 21.7% guanine (G). In all four observed genotypes, the adenine (A) nucleotide has the most percentage, while cytosine (C) the least. The multiple sequence alignment of the deduced MSTN nucleotide sequences of animals from four different genotypes are presented in Figure 2. Based on the results, all genotypes were lacking a late segment between 380-bp and 417-bp of intron 1 of MSTN gene.
All sequences match thoroughly for intron I except for three regions showing single nucleotide polymorphism (SNP). These three regions were in 224bp, 226bp and 242bp situations, following three substitutions as c.224C>T; c.226A>G; c.242G>T. All three SNPs were located about the same region. One substitution (c.224C>T) was in intron 1 and resulted in an amino acid substitution of serine (Ser) with phenylalanine (Phe) at codon 224; another substitution (c.226A>G) replaced amino acid of methionine (Met) with valine (Val); and the third (c.242G>T) replaced glycine (Gly) with valine (Val). The entropy plot of nucleotide sequences is shown in Figure 3.

A phylogenetic tree based on nucleotide sequences (excluding intron I) was generated using Neighbor-Joining method in MEGA version 6.0 (Tamura et al., 2007). Phylogenetic analysis of Makuei sheep breeds and their linkage with foreign sheep breeds is shown in Figure 4.
Molecular characterization of MSTN gene in Makuei Sheep

Figure 3. Entropy plot of the MSTN gene nucleotide showing the divergent sites (as peak) in Iranian Makuei sheep breed. A score of 0 stands for complete conservation, whereas higher positive values indicate greater diversity between the sequences.

Figure 4. Phylogenetic analysis based on deduced nucleotide sequences of MSTN for Iranian Makuei sheep breed with other sheep of different origin registered to Genebank.

The result of phylogeny tree branched into nearly two different set of groups, with Dorset Down, Merino, Suffolk and Perendale grouped together separate to the rest of the animals; their difference lies with the locations of mutations and type of mutations. The results also revealed that the Iranian native Makuei sheep breed standing in the middle of the tree has more sequence identity at nucleotide level with Poll Dorset breed than it has with others.
Discussion

PCR amplification of intron I of MSTN gene and SSCP analysis

The naturally occurring mutations in MSTN gene have been associated with hypertrophy in cattle (Grobet et al., 1997; Kambadur et al., 1997; McPherron and Lee, 1997; Karim et al., 2000), human (Ferrell et al., 1999) and sheep (Boman et al., 2011). In the present study, we investigated variation of the intron 1 region of MSTN gene in Iranian Makuei sheep breed using SSCP method. The results of this study are in agreement with the results obtained by Hickford et al. (2010), who identified five alleles in the MSTN gene in both New Zealand and Makuei sheep breeds. Likewise, Ansari et al. (2011) reported three SSCP patterns in intron 1 region of Baluchi sheep. Other researches, using PCR products amplified from MSTN gene, from a total sample of 60 sheep, including 9 Chinese indigenous sheep breeds and 1 imported sheep breed, identified SNPs in a 378-bp fragment including intron II and exon III of the MSTN gene (Gong et al., 2009).

Association of intron I of MSTN gene polymorphism with biometric traits

The most number of animals in this study showed the TAT genotypes, and only these showed significantly different Heart Girth (HG) and Leg Circumference (LC) compared with those lacking thereof (Table 1).

Since substitution of c.226A>G leads to replacement of methionine (Met) with valine (Val), we came to the conclusion that nucleotide “A” exerted a greater influence on the difference related to HG and LC traits.

However, because the polymorphism lies in non-coding DNA, drawing conclusions as to whether this genetic variation affects MSTN gene activity did not seem to be straightforward. It may affect mRNA splicing or be linked to variation elsewhere in the coding sequence, which subsequently affects the amino acid sequence.

Other relationships between the MSTN gene and growth have been examined. The effect of genetic variation in the MSTN gene on growth and carcass traits has been investigated in 517 male Romney lambs from 17 sire-lines born on a South Island New Zealand farm (Hickford et al., 2010). A general linear mixed model revealed that the presence of allele A in lambs is associated with decreases in leg, loin, and total yield of lean meat, whereas the presence of allele B is associated with increases in and proportion of loin yield (Hickford et al., 2010).

Zandi et al. (2013) investigated the polymorphism of promoter and exon 1 region of myostatin gene in native chickens of West Azerbaijan. They reported three different genotypes (AA, AB and AC). In their study, the SSCP pattern did not have any significant effect on phenotypic value of different traits, but the breeding value of the body weight at 12 weeks was significantly affected by SSCP pattern.

Several mutations in the MSTN gene may be related to various phenotypes in cattle breeds (Grobet et al., 1997; Kambadur et al., 1997; Gill et al., 2009). Li et al. (2002) have identified mutations in exons II and III of the swine MSTN gene. Zheng et al. (2008) have found two SNPs in the promoter region of the MSTN gene (T769G, C543T) and one in intron I (A1632G) in 10 goose breeds. In double-muscled cattle, seven mutations occur in the MSTN gene, namely, nt821 (del 11), nt419
Four novel SNPs (G2283A, C7552T, C7638T, and T7661A) were detected in exons of the myostatin gene in the Bian chicken. The findings from the least square means showed that Bian chickens with EE and DE genotypes had significantly higher body weight, at 6-18 weeks of age, than those of the DD genotype (Zhang et al., 2011).

In this research, three novel SNPs were found in the intron I that have not previously been reported in any sheep breed. Among these SNPs, the one in c.226A>G location showed significant difference in two biometric traits (HW: height at withers and LC: leg circumference). Thus it could be concluded that this substitution (c.226A>G) in TAT genotypes may control the biometric traits. However, more studies in more breeds are needed to validate the claim.

**Analysis of Intron I region of MSTN gene sequence**

In the nucleotide sequence level, MSTN sequences of Iranian Makuei sheep breed exhibited similarity among themselves except three regions (224–226 and 242bp position). Analysis of entropy plot data suggested the presence of three divergent sites within the highly conserved parts of Iranian Makuei sheep MSTN nucleotide sequences (Figure 2). The highly conserved sites are 1–221bp, 227–241bp and 243–417bp.

In this study, the PCR-SSCP method enabled the researchers to identify previously unexplained variation in intron 1 of MSTN. SNP variations lead to amino acid substitution of Ser with Phe, Met with Val and Gly with Val in the intron I locus.

In the previous studies, researchers reported some variations in whole MSTN gene such as c.-1129C>T (Clop et al., 2006) in the promoter region; c.-41A>C, c.-38C>T and c.-31delT (Du et al., 2005; Clop et al., 2006; Kijas et al., 2007; Gan et al., 2008; Zhou et al., 2008; Boman et al., 2009; Boman and Vage, 2009; Sjakste et al., 2011) in the 5’UTR; c.101G>A (Zhou et al., 2008) in exon 1 region; c.373+18 T>G, c.373+241C>T, c.373+243A>G, c.373+246C>T, c.373+249C>T, c.373+259T>G and c.373+323T>C (Clop et al., 2006; Kijas et al., 2007; Gan et al., 2008; Zhou et al., 2008; Hickford et al., 2010; Sjakste et al., 2011) in intron 1 region; c.747+164G>A, c.748-810T>C, c.748-575C>A and c.748-568T>C (Clop et al., 2006; Kijas et al., 2007; Gan et al., 2008). Our results partially match those of previously published studies, the only difference being the location of the SNPs.

The variation described by Zhou et al. (2008) led to an important amino acid substitution of Glu with Gly at the codon 34. The results of Zhou et al. (2008) in NZ Romney sheep and Merino, Corriedale with crossbred sheep and here in Makuei sheep breed similar to the NZ Romney, dismiss any claim concerning occurrence of any genetic variation in the MSTN coding site (Clop et al., 2006; Kijas et al., 2007).

Genetic variation present in intronic regions may affect mRNA splicing and subsequently affect the amino acid sequence produced from the transcript, but not the coding region. If it had any effect on the operation of the donor splice-site, it could have influenced pre-mRNA secondary structure during skeletal muscle development.
(Sjakste et al., 2011) and ultimately MSTN function. Confirmation requires further analysis of MSTN expression in sheep carrying this substitution.

The three novel substitutions reported here, together with the substitutions observed previously (Du et al., 2005; Clop et al., 2006; Kijas et al., 2007; Gan et al., 2008; Hadjipavlou et al., 2008; Zhou et al., 2008; Boman et al., 2009; Boman and Vage, 2009; Hickford et al., 2010; Sjakste et al., 2011) suggest that ovine MSTN is highly variable and, therefore, further potential variations may exist in MSTN.

The variations may affect MSTN activity and, consequently, muscle growth and biometric traits. It might sound reasonable to expect that new variations unique to that breed, or at the very least uncommon in other breeds may come to light, as further breeds are studied. It can also be suggested that close relationship between Iranian Makuei sheep breed and the same geographical ecoclimatic zone might influence Poll Dorset.

**Conclusions**

In this study, the variation of intron I region of MSTN gene in one of the Iranian native sheep (Makuei sheep) and its relationships with BVs of biometric traits are reported. Intron 1 region of MSTN sequences was found to be very similar to Poll Dorset sheep compared to other foreign sheep. Role of these polymorphisms in the MSTN activity in sheep, specifically in the meat breeds has yet to be analyzed. MSTN, thus, may be one of the several genes that regulate changes in skeletal muscle and body mass, as body mass affects the biometric traits.

Breeding programs in most Iranian sheep research centers are solely based on phenotypic characters. The current study confirmed the importance of molecular studies along with phenotypic data for improving the further genetic gain.

Additional researches are needed to characterize the completely ovine MSTN gene variation across an extended region of the gene and in a large variety of sheep breeds from around the world.

In summary, the work presented in this study suggests that ovine MSTN is diverse and requires further characterization across various breeds. Genetic variation in MSTN gene can be used as a selection tool for improving carcass traits and, possibly, more meat production may be feasible. However, future studies to investigate the effect of specific MSTN variants on sheep meat production and meat quality traits and in different sheep breeds are still required.

**Acknowledgment**

The Department of Animal Science and the Institute of Biotechnology, Urmia University, Urmia, Iran supported this research work.

**References**

Molecular characterization of MSTN gene in Makuei Sheep


M. Farhadian and A. Hashemi


Received: 16 VIII 2015
Accepted: 2 III 2016