Novel SNPs of the ABCG2 Gene and Their Associations with Milk Production Traits in Iranian Holstein Bulls

S. A. Mousavizadeh¹, A. Salehi¹∗, M. Aminafshar², M. B. Sayyadnejad³, and M. H. Nazemshirazi⁴

ABSTRACT

ABCG2 (ATP binding cassette subfamily G member 2) gene, located on chromosome 6 encodes the ABCG2 protein that transports various xenobiotics, cytostatic drugs across the plasma membrane as well as cholesterol into milk. A single nucleotide change (A/C) in base 86 of exon 14 is capable of encoding a substitution of tyrosine with serine in the ABCG2 gene and increase milk yield while decreasing milk fat and protein concentrations. The major aim followed in this research was to study Single Nucleotide Polymorphisms (SNPs) of ABCG2 gene and their association with milk production traits in Iranian Holstein bulls. Genomic DNA of 105 bulls was extracted from semen samples using highly Pure PCR template preparation kit. Primers were designed through Oligo software (Version 5.0) and utilized in PCR. Then the PCR fragments were sequenced. The A/C substitution in base number 86 of exon 14 was observed with 2% frequency which affected protein percentage (P< 0.05). Some SNPs were detected for the first time in intron 13, exon and intron 14 in comparison with sequences in the NCBI database. A deletion mutation in base number 20 (T/−) and a missense mutation in base number 67 (A/G) of exon 14 that cause the substitution of serine with glycine were discovered which were significantly associated with protein yield and fat percentage, respectively (P< 0.05). Furthermore, significant association was observed between fat percentage and mutations in base numbers 4,133 (T/C) and 4,137 (T/G) of intron 13 (P< 0.05). Substitutions in base numbers 2 (T/C) and 55 (G/C) of intron 14 resulted in a significant effect on fat yield and fat percentage (P< 0.05).

Keywords: ABCG2 gene, Holstein bulls, Milk production traits, Polymorphism.

INTRODUCTION

Various studies have provided candidate genes as based on their physiological roles in different traits and diseases. Research of Quantitative Trait Loci (QTL) on chromosome 6 (BTA6) proposed such genes as PPARGC1A, PKD2, SPP1, OPN and ABCG2 as candidate genes which affect milk components (Cohen et al., 2004; Olsen et al., 2005; Weikard et al., 2005; Olsen et al., 2007).

Many QTL studies demonstrated the effects of the PPARGC1A gene on milk production also because of the participation of this gene in fat metabolism (Weikard et al., 2005). Khatib et al. (2007) studied associations between transverse mutation A/C at position 3,359 of PPARGC1A gene and milk production traits in two Holstein cattle populations. They concluded that A

¹ Department of Animal and Poultry Science, Abureyhan Campus, University of Tehran, Pakdasht, Islamic Republic of Iran.
² Corresponding author, email: arsalehi@ut.ac.ir
³ Department of Animal Science, Faculty of Agriculture and Natural Resources, Science and Research Branch, Islamic Azad University, Tehran, Islamic Republic of Iran.
⁴ Animal Breeding Center of Iran, Karaj, Islamic Republic of Iran.
⁵ Department of Molecular, Iran Central Veterinary Laboratory, Tehran, Islamic Republic of Iran.
allele was associated with significant positive effects on protein percentage and as well a decrease in milk yield in both populations. Olsen et al. (2005) proposed PKD2 gene (polycystin 2) as a candidate gene in dairy cattle. Since calcium is the major osmotic constituent in milk, variations in PKD2 expression could affect the water content of the milk, and circuitously increase milk fat and protein percentages (Olsen et al., 2005). He suggested three genes (ABCG2, OPN and PKD2) as candidate genes for the QTL.

Cohen et al. (2004) suggested that SPP1 plays a crucial role in mammary gland differentiation and branching of the mammary epithelial ductal system. The ABCG2 and SPP1 were expressed in the mammary gland of bovine and increased from parturition through lactation (Cohen-Zinder et al., 2005). Results of research have revealed significant associations between OPN gene polymorphism and milk quantitative traits (Leonard et al., 2005; Schnabel et al., 2005; Khatib et al., 2007; Olsen et al., 2007).

ABCG2 gene is a member of ATP-binding cassette transporters which are revealed to deliver xenobiotics and metabolites across the canalicula space into the bile (Leslie et al., 2005). Jonker et al. (2005) reported that ABCG2 was not expressed in virgin mice but was strongly explicated during late pregnancy and especially through lactation. They demonstrated that ABCG2 expression is responsible for the active secretion of such clinically and toxicologically important substrates as dietary carcinogen PhIP, anticancer drug Topotecan and antiulcerative Cimetidine into mouse milk. The ABCG2 gene seems to play a key role in stem cell regulation and also in hypoxic defense mechanisms in human (Zhou et al., 2001; Sarkadi et al., 2004). ABCG2 expression is significantly enhanced during lactation and is accountable for the secretion of vitamin K3 or cholesterol into milk (Van Herwaarden et al., 2007; Farke et al., 2008).

Several SNPs have been identified in the ABCG2 gene which are of the most significant associations with milk yield. Fat and protein concentrations were demonstrated for A to C substitution in base number 86 of exon 14, causing a change of the amino acid from tyrosine to serine, at location 581 (Y581S). The mutation was associated with an increased milk yield as well as decreased fat and protein percentages on Bos Taurus (Olsen et al., 2005; Cohen-Zinder et al., 2005; Olsen et al., 2007).

Holstein registered heifers from Europe, the United States, and Canada during were imported 1970s and early 80s to establish the intensive dairy cattle husbandry in Iran. More than 90% of milk presently sold on the free market is Holstein cows’ milk. Registered Holstein dairy cow population amounts to about 1 million representing 12.5% of the total cattle population nationwide.

The ultimate aim followed in this study was to detect Single Nucleotide Polymorphism in the ABCG2 gene as a candidate gene in Iranian Holstein registered bulls and to detect the influence of polymorphisms on milk production traits.

MATERIALS AND METHODS

DNA Extraction and SNP Genotyping

Semen samples were obtained from 105 Iranian Holstein bulls representing resource populations. Genomic DNA was extracted using high Pure PCR template preparation kit (Roche Company kit, CAD No= 11796828001). The quantity and quality of the extracted DNA were assessed through spectrophotometry and electrophoresis on 2 percent Agarose gel. To amplify 240 bp region including partial sequences of intron 13 (base number 4,060 to 4,141), the entire of exon 14 and partial sequences of intron 14 (base number 1 to 68) of ABCG2 gene, primers were designed through Oligo software (version 5.0). Forward and reverse primers were detected as 5’-GTATTACGGAGACTGTCAGGG-3’ and 5’-
GGCTTTATTCTGGCTGTTTCC-3`

respectively. The PCR amplification was adjusted in the best possible condition, 5 µl (150 ng µl⁻¹) of DNA samples being added to 20 µl of PCR mixtures containing 5 µl PCR buffer (10X), 0.5 µl MgCl₂ (1.5 mM), 1 µl of each dNTPs (10 mM), 1 µl of each primer (10 Pmol µl⁻¹) and 0.5µl of Taq DNA polymerase for (5 unit µl⁻¹). Amplification reactions were conducted in a thermal cycler with an initial denaturation at 95°C for 15 minutes and 35 cycles at 95°C for 30 seconds, 50°C for 30 seconds, and 72°C for 40 seconds followed by a final extension step at 72°C for 5 minutes.

PCR fragments were purified by means of Qiaquick PCR purification kit (Qiagen Company kit, CAD No= 28104) and then sequenced through ABI 3730 XL 16 Capillary Sequencer and ABI 3730 XL 96 Capillary Sequencer applying Sanger and Dideoxy Chain Termination methods respectively. Eventually, results of sequencing were aligned, making use of Blast software (NCBI) to find similarities vs. differences with the submitted sequence of the ABCG2 gene in the NCBI database.

**Statistical Analysis**

Breeding values of quantitative traits were estimated through animal model. Phenotypic records were assessed at first lactation. The records of 305 days and twice milking per day were applied. Only records of cows with calving between 18 and 38 months of age and exceeding 90 days of record were included while animals of no records being excluded. Analyzed traits were Milk Yield (MY), milk Fat Yield (FY), Fat Percentage (FP), milk Protein Yield (PY) and Protein Percentage (PP). Restricted Maximum likelihood Method (REML) based on average information algorithm using ASRMEL programs (version 3.1) was employed to estimate the variance of components (model number 1). The data file included 105 Iran-born Holstein bulls, born between 1983 and 2002. The EBVs were based on phenotypic records of daughters including 240,000 to 450,000 data items, varying from 59 to 3,632 per bull scattered within 10 to 1,074 herds, recorded from 1967 to 2008.

Model 1) \( Y_{ij}= HYS_j+ Age_i+ A_i+ e_{ij} \)

Where, \( Y_{ij} \) is observation related to favorite traits of \( i^{th} \) animal, \( HYS_j \) the \( j^{th} \) herd-year season combinations effect (as fixed effect), \( Age_i \) age at first calving (as a covariate) for \( i^{th} \) animal, \( A_i \) representing the additive genetic effects of \( i^{th} \) animal and \( e_{ij} \) standing for random residual effect.

Generalized Linear Model (GLM) was made use of to examine associations between milk quantitative traits and the detected ABCG2 SNPs using model number 2.

Model 2) \( Y_{ij}= \mu + SNP_i + e_{ij} \)

In which \( Y_{ij} \) is the predicted breeding value for milk production traits related to the \( i^{th} \) bull and \( j^{th} \) SNP. \( SNP_i \) is the \( i^{th} \) SNP and \( e_{ij} \) the residual random effect. The allele substitution effects were evaluated by regressing EBVs on the number of copies of each allele carried by each animal with the results being presented as the mean±standard deviations.

**RESULTS**

The length of PCR products was equal to 240 bp. Results of sequencing indicated new SNPs in comparison with the recorded sequence of the ABCG2 gene in the NCBI database (Accession number: AJ871176). All the new polymorphisms were submitted in the NCBI and got the accession number. Statistical results demonstrated significant associations between new mutations in intron 13, exon plus intron 14 and the breeding value of milk quantitative traits.

In intron 13, the \( T/C \) mutation (base number 4133, GeneBank accession number: JQ398809) with 7 percent frequency and \( T/G \) mutation (base number 4137, GeneBank accession number: JQ398800) with 4 percent frequency showed their significant effects on fat percentage (P< 0.05).
A deletion mutation in base number 20 (T/-, Gene Bank accession number: JN811066) with a frequency of 4 percent along with a missense mutation in base number 67 (A/G, Gene Bank accession number: HQ730358) of 8 percent frequency causing the substitution of serine to glycine in exon 14 exerted a significant effect on protein yield and fat percentage, respectively (P<0.05).

A/C mutation (GeneBank accession number: JQ398810) in base number 86 of exon 14 (Y581S) was observed in the Holstein bull population of Iran. Frequencies of allele A and C were recorded as 0.98 and 0.02, respectively. The non-conservative Y581S mutation in ABCG2 affected protein percentage (P<0.05).

In addition, the results revealed that polymorphisms in base numbers 2 (T/C, GeneBank accession number: JQ398798), and 55 (G/C, GeneBank accession number: JQ398814) in intron 14 with frequencies 13 and 7 percent, respectively had significant influence over fat yield and fat percentage (P<0.05).

The average breeding values of milk quantitative traits were evaluated for wild and mutant alleles of ABCG2 gene. Results revealed that, the average breeding values of protein percentage, fat yield and fat percentage were higher for ABCG2A as compared with ABCG2C (base 86 of exon 14), while being lower in milk and protein yields (Table 1).

### DISCUSSION

A main goal of dairy cattle genomic research is to identify genes underlying the variation of milk production traits that can be functional in breeding programs. The candidate gene approach, is to supply tools for studying causative SNPs that influence milk components. Genes with a major effect on milk quantitative traits in dairy cattle could be involved in various physiological pathways including triglyceride synthesis diacylglycerol acyltransferase 1 (DGAT1) and ATP binding cassette subfamily G member 2 transporters (ABCG2).

The SNP showing significant association with milk components would afford a main opportunity for Marker-Assisted Selection (MAS) programs in livestock (Khatib et al., 2007). A study of the substitutions in ABCG2 gene alleles and their effects on

### Table 1. Estimated average breeding value of milk quantitative traits for SNPs in the ABCG2 gene (±SE).

<table>
<thead>
<tr>
<th>Gene Region</th>
<th>Allele</th>
<th>MY ²</th>
<th>FY ²</th>
<th>FP ²</th>
<th>PY ²</th>
<th>PP ²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exon 14 ABCG2A</td>
<td>-160.1(±50.3)</td>
<td>-1.9(±1.32)</td>
<td>0.002(±0.02)</td>
<td>1.9(±1.75)</td>
<td>-0.009(±0.01)</td>
<td></td>
</tr>
<tr>
<td>Base 86 ABCG2C</td>
<td>294.5(±287.56)</td>
<td>-2.3(±3.4)</td>
<td>0.008(±0.07)</td>
<td>-1.08(±1.81)</td>
<td>0.009(±0.03)</td>
<td></td>
</tr>
<tr>
<td>Intron 13 ABCG2A</td>
<td>-165(±122.32)</td>
<td>-1.68(±6.26)</td>
<td>0.006(±0.11)</td>
<td>1.44(±10.54)</td>
<td>-0.01(±0.04)</td>
<td></td>
</tr>
<tr>
<td>Intron 13 ABCG2C</td>
<td>-162.2(±219.39)</td>
<td>-1.77(±11.55)</td>
<td>0.006(±0.11)</td>
<td>1.44(±10.54)</td>
<td>-0.01(±0.04)</td>
<td></td>
</tr>
<tr>
<td>Intron 14 ABCG2G</td>
<td>-5.8(±5.09)</td>
<td>-6.38(±1.2)</td>
<td>0.017(±0.01)</td>
<td>7.24(±1.75)</td>
<td>-0.03(±0.01)</td>
<td></td>
</tr>
<tr>
<td>Intron 14 ABCG2C</td>
<td>-193.3(±52.61)</td>
<td>-1.3(±1.34)</td>
<td>0.007(±0.08)</td>
<td>1.67(±3.92)</td>
<td>-0.009(±0.03)</td>
<td></td>
</tr>
</tbody>
</table>

² Milk Yield; ² Milk Fat Yield; ² Fat Percentage; ² Milk Protein Yield, ² Protein Percentage.
New Novel SNPs of the ABCG2 Gene

milk components are economically desirable for most selection indexes used in dairy cattle breeding programs. Through which rates of genetic gain could be promoted by direct selection on the alleles. 

ABCG2 encoded a protein that is a member of the ATP binding cassette super family (ABC). A wide variety of drugs and various xenobiotics are extruded through the protein and across the plasma membrane (Litman et al., 2000). A/C substitution in exon 14, is capable of encoding substitution of tyrosine to serine in the ABCG2 gene and affect milk quantitative traits (Cohen-Zinder et al., 2005; Olsen et al., 2007). Ron et al. (2006) reported allele frequency of A/C mutation (Y581S) in 35 breeds. They proposed that the allele A of the ABCG2 gene was predominant in all the populations. The recognition of allele C (base 86 of exon 14) only in Bos taurus breeds may designate that allele A of ABCG2 is the ancestral and the Y581S the substitution that occurred following a separation of Bos indicus from Bos taurus genealogy over 200,000 years ago (Ron et al., 2006).

Here, the emphasis is mainly upon the ABCG2 gene polymorphisms as the most functional candidate gene affecting milk traits in Iranian Holstein bulls. An A/C mutation in base 86 of exon 14 was observed with 2% frequency, allele C had rare frequency in Iranian Holstein bulls similar to the other breeds. Statistical results also revealed the same influence of A/C polymorphism on milk quantitative traits match with the results obtained by other researchers (Cohen-Zinder et al., 2005; Olsen et al., 2007; Komisarek and Dorynek, 2009). New SNPs have been identified in the ABCG2 gene which are associated with milk quantitative traits. Moreover, the results of the present study revealed that SNPs in ABCG2 are in association with milk protein and protein percentage traits. In the era of a wide variety of SNP genotypes in animals, it is objective to report estimates out of an optimally fitted model. To be able to recognize such a model, the suitable model must be fit to be used among the theoretical possible models.

According to the results of the present study, utilization of ABCG2 gene variations as an index in the Gene Assisted Optimization Index is proposed beside use of other genes affecting traits (such as multiple genes) that are the leading cause of improvement in genetic gain within the domain of animal selection. This study would be functional in Holsteins’ breeding programs in Iran. More researches are suggested to study the ABCG2 polymorphisms as a candidate gene along with its relationships with economical traits, due to its chromosomal position as well as its key role in milk productions.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the Animal Breeding Centre of Iran (Karaj, Iran) and also Molecular Department Central Veterinary Laboratory for providing semen samples and for their other generous contributions. Special thanks are extended to Dr Davood Kolbehdari, Dr Mirtorabi, Dr Rouhollah Abdollahi and Mrs. Zahra Namvar for their sincere assistance.

REFERENCES


جندشکلی های جدید زن ABCG2

زن ههلشماتین ایرانی

س. ع. موسوی زاده، ع. صالحی، م. امین افشار، م. ب. صیاد نژاد، م. ح. ناظم

چکیده

زن ABCG2 (ATP binding cassette sub family G member 2) گروه ۲ مارکر شناسایی راوی کروموزوم شماره ۶ گاو بیان پروتئین ABCG2 در انتقال مواد داروئی از غشاء پلاسمای کلسترول به شیر نقض دارد. در اثر جهش در چند شماره ۸۴ و ۸۵ در شماره ۷۵، آل ۱۴ آثل A به آثل B تبیین شده و در اثر آن اسید آمیونی بروز کرده بی‌سیرین تغییر می‌یابد که افزایش حساسیت به داروهای چربی و پروتئین اemies (GENES) و ارتباط آن با صفات ABCG2 تولیدی شیر در گاوگان نزدیک ههلشماتین ایرانی تولد نمود داده. DNA تولیدی همه آن نموده اسیرم گاوگانی از ۱۰۵ نمونه اسیرم گاوگانی است. استخراج شدن. آغازگرها توسط نرم‌افزار Oligo TCR (مدل ۵) طراحی و برای تکثیر قطعات مورد نظر استفاده گردید. پس از تعیین تولید محصولات (PCR) SNPs (Single Nucleotide Polymorphisms) اولین مرتبه در این مطالعه در چندین از ۱۲۰ پروتئین و اینترن در مقایسه با توالی موجود در NCBI شناسایی شدند. جهش حذفی در چند شماره ۲۰ تا (T/G) (A/G) ۱۴ در اثر اسید آمیونی سرین به اسید آمیونی گلیزین تبیین می‌شود مشاهده شد که به ترتیب با مقدار پروتئین و درصد چربی ارتباط قابل توجهی داشتند (P<0.05). همچنین ارتباط معنی داری بین جهش با چربی شماره ۴۳۳ ۱۳۷ تا (T/C) و ۱۳۱ تا (T/C) ۹۴ اثر فعالیت در پی در بار شماره ۲ (G/C) ۵۵ اثر قابل توجهی بر روی پیامد و درصد چربی (P<0.05).