Meta Gene 6 (2015) 85–90

Contents lists available at ScienceDirect

Meta Gene

Identification of functional consequence of a novel selection signature in *CYP11b1* gene for milk fat content in *Bubalus bubalis*

J. Maryam [a,*], Masroor Ellahi Babar [b], Asif Nadeem [a], Tahir Yaqub [c], Abu Saeed Hashmi [a]

[a] Institute of Biochemistry & Biotechnology, University of Veterinary & Animal Sciences, Lahore 54000, Pakistan
[b] Virtual Universities, Lahore 54000, Pakistan
[c] Department of Microbiology, University of Veterinary & Animal Sciences, Lahore 54000, Pakistan

ARTICLE INFO

Article history:
Received 1 January 2015
Revised 1 September 2015
Accepted 8 September 2015
Available online 22 October 2015

Keywords:
SNPs
Milk fat %age
River buffalo
CYP11b1 gene
Protein model
CAPS

ABSTRACT

Genomic selection for traits of economic importance is an emerging approach carrying tremendous potentials. Many of polygenic traits as milk fat, protein and yield have been characterize at genomic level and important selection signatures have been identified. Cytochrome P450 enzymes are potential loci for affecting many of dairy capabilities. Present study was conducted for genomic dissection of *CYP11b1* gene in riverine buffaloes and seven genetic variations were identified. Out of these, one novel polymorphism (p.A313T) was found well associated with milk fat %age. AB genotyped buffaloes were found to have higher milk fat %age (8.9%) for this loci. p.A313T was further validated at larger data set by restriction digestion using CviAII enzyme. Functional consequences of this locus were also predicted by studying three dimensional structure of *CYP11b1* protein. For this purpose, 3D protein model was predicted by homology modeling, secondary structural attributes were determined, signal peptide was predicted and a transmembrane helix was also identified. One of polymorphism (p.Y205L) was found in the vicinity of functionally significant F-G loop region, which is the part of protein gets attached to the inner mitochondrial membrane. But this variation could not be associated and needs further investigation. p.A30V, a popular selection marker in cattle, was found in buffaloes as well but could not be associated and might need further confirmation on larger data set. Results of this study illustrate the impending potential of this gene in determining dairy capabilities of buffaloes and might have a role in selection of superior dairy buffaloes.

1. Introduction

Among species of livestock, buffalo stands out as an efficient converter of poor quality roughages into highly valuable products as milk and meat. As a major contributed in overall milk (68%) and meat (23%) production of the region, buffalo breeds of Pakistan are admired internationally as well (Afzal, 2010). Dairy potentials of these animals have not been explored much at genetic level. Majority of studies endorse existing genomic variants identified in cattle, which are not very much helpful in buffalo selection. Although buffalo-cattle genome homology is quite higher (more than 85%) but there have been circumstantial variations at phenotypic level which are needed to be evaluated at molecular level. Fat is the major buffalo milk preferential, which has been noted up to 8% or even in some of individuals has been found more than 12% (Bilal and Sajid 2005). This variation strengthens the idea of genetic basis of this trait and paves the way towards identification of genetic basis of milk fat content that is ultimate resolution to the improved milk quality.

The Cytochrome P450 enzymes are involved in steroid hormone biosynthesis. These membrane-bound proteins associated with either the mitochondrial membranes (*CYP11b1* and CYP11b2) or the endoplasmic reticulum (CYP17, CYP19, and CYP21) (Payne and Hales 2004; Brettes and Mathelin 2008). Many of previous reports provide information about the association of *CYP11b1* gene with milk quality traits especially fat content. Present study was planned to genomically dissect *CYP11b1* to identify novel variations in Pakistani buffaloes.

Exonic part of the *CYP11b1* gene was sequenced by Sanger's method of DNA sequencing. Seven polymorphisms were identified and were statistically analyzed by calculating Chi^2 to study Hardy Weinberg Equilibrium (HWE) ($P < 0.05$). Variations obeying HWE were selected for association testing and only one out of seven polymorphisms was associated with milk fat% age. Finally, 3D structure of *CYP11b1* was predicted to locate this associated functional variant by bioinformatics

* Corresponding author.
E-mail address: maryam.javed@uvas.edu.pk (J. Maryam).

http://dx.doi.org/10.1016/j.mgene.2015.09.002

Table 1
Primer pairs designed to amplify exonic regions of *CYP11b1* gene.

Primer name	5′–3′ Sequence	Product size
CYP F1	AGGCTTCCTGGTCTGG	497
CYP R1	CCCTCCCTACCCCTTT	
CYP F2	AGCACAGAACGCAGACC	457
CYP R2	AGGAAGTCGAGCCCTTG	
CYP F3	TGTCTGGCTGGTTTCACT	400
CYP R3	ACTAAGGTGCTGGCTGTG	
CYP F4	GGAGTCTGACCCTGGACATC	461
CYP R4	ACCATAACGAAGCCACAAGC	
CYP F5	AGGACGTGGAGAATTGG	376
CYP R5	AGCTGGAGGCATAGATTG	
CYP F6	CGGAGTGGAGGGACATGG	300
CYP R6	GTCAAGCCCAGCAAGAGG	
CYP F7	TAGCAGCAGTAGCAGCAGGA	563
CYP R7	GACAGAGGCAGGGTTCCAC	
CYP F8	CTGCACCATGTGAGTGG	335
CYP R8	GGGCCTGTAGGAGAAAGA	

Table 2
Polymorphic sites detected in the *CYP11b1* gene.

SNPs	Wild type	Mutation	Transition/tranversion	Amino acid substitution
p.A30V	G	A	Transition	Alanine to valine
p.55589	A	C	Transversion	Intronic
p.M111R	G	A	Transition	Methionine to arginine
p.Y205L	T	G	Transversion	Tyrosine to leucine
p.T300N	A	G	Transition	Threonine to asparagine
p.T312M	G	T	Transversion	Threonine to methionine
p.A313T	C	A	Transversion	Alanine to threonine

software to evaluate the functional consequences of this novel variation. The results of this research provide baseline information in our hunt for genetic signatures controlling economically important traits to improve the milk quality and productivity of indigenous buffalo breeds.

Table 3
Allelic frequency and HWE of identified variants in *CYP11b1*. ($P < 0.05$).

SNP ID	Allele frequency		Chi^2 ($P<0.05$)
p.A30V	A	B	0.134801**
	0.3415	0.6585	
p.55589	A	B	0.0010*
	0.6585	0.3415	
p.M111R	A	B	0.000083*
	0.6707	0.3293	
p.Y205L	A	B	0.1703**
	0.3780	0.6225	
p.T300N	A	B	0.00014*
	0.3659	0.6341	
p.T312M	A	B	0.06812**
	0.6829	0.3171	
p.A313T	A	B	0.17016**
	0.6029	0.3971	

* Significant.
** Non-significant.

2. Materials and methods

2.1. Sampling strategy.

Nili–Ravi buffaloes were selected at the first month of second lactation and blood and milk sampling was conducted. Animals were selected on the basis of milk fat %age and two groups were constructed. In group-A, animals (n = 35) with butter aft %age more than 8% were included and in group-B, animals (n = 35) with butter fat less than 8% were added. A total of 70 animals were included in milk and blood sampling. Many of Govt. livestock farms were visited for animal selection and sampling as Buffalo research institute (BRI), Pattoki, Livestock experimental station (LES) Bahadurnagar and Buffalo colony, Karachi.

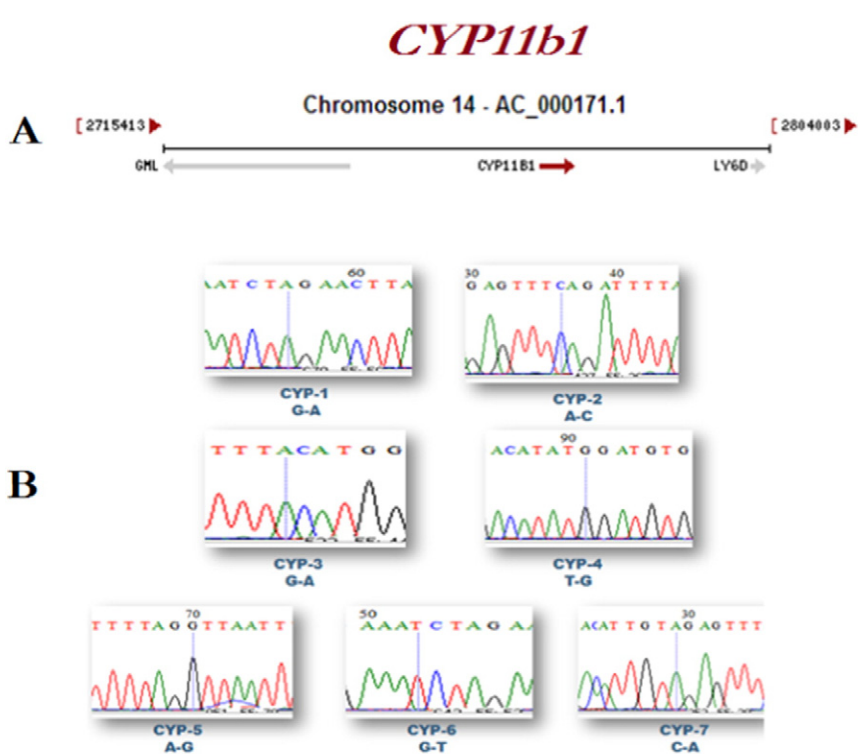

Fig 1. Genetic organization of *CYP11b1* gene illustrates DNA sequence variations. A: Chromosomal location of *CYP11b1*. B: Seven genetic variants identified in *CYP11b1* gene.

Table 4
Genotypic frequency for all loci of *CYP11b1*.

0.2683	0.1463	0.5854
0.5610	0.1951	0.2439
0.5806	0.0645	0.3548
0.5122	0.2195	0.2683
0.2683	0.1951	0.5366
0.5484	0.0645	0.3871
0.1707	0.5854	0.2439

Table 5
Single marker association by one way ANOVA (P < 0.05).

Genetic variations	AA (Mean ± SE)	AB (Mean ± SE)	BB (Mean ± SE)	P-value (P < 0.05)
p.A30V	n = 26	n = 14	n = 58	0.253261
	7.58 ± 0.1497	7.05 ± 0.7136	8.88 ± 1.0575	
p.55589	n = 56	n = 19	n = 24	0.469714
	7.26 ± 0.8841	6.3 ± 0.8832	8.48 ± 1.557	
p.M111R	n = 17	n = 58	n = 24	0.060839
	6.06 ± 0.6765	8.3 ± 1.4989	5.68 ± 0.9521	
p.Y205L	n = 51	n = 21	n = 26	0.006907
	4.7 ± 0.7627^b	4.9 ± 0.3317^b	7.94 ± 0.7833^a	
p.T300N	n = 26	n = 19	n = 53	0.306324
	6.86 ± 0.3059	7.1 ± 1.2275	8.52 ± 0.8027	
p.T312M	n = 54	n = 06	n = 38	0.218845
	6.86 ± 1.1374	5.6 ± 0.6758	8.92 ± 0.6003	
p.A313T	n = 06	n = 58	n = 35	<.0001
	5 ± 0.3536^b	8.9 ± 0.2915^a	4.9 ± 0.3317^b	

Notes: P-value refers to the results of association analysis between each SNP and milk Fat %age. Means within a row with different superscripts differ (P<0.05).

2.2. Genomic DNA amplification and sequencing.

DNA was extracted by organic method reported by Maryam et al. (2012). Pooling of individuals of group-A and B was performed separately and these pools were used for amplification of exonic region of *CYP11b1* gene by using specific sets of primers designed by using Primer3 (primer3.ut.ee) (Table 1). Amplification was performed by using 0.5 μl of each primer, 2.5 μl 109 PCR buffer, 2.5 mM each of dNTP, and 1 U of Taq DNA polymerase. Then, each PCR product was sequenced by Sanger's chain termination method using the ABI3730XL (Applied Biosystems, Foster City, CA). Sequences from both groups were aligned on BLAST resource of NCBI and a total of seven variations were identified (Table 2, Fig. 1).

2.3. Statistical analysis

POPGENE (http://www.ualberta.ca/fyeh/) version 1.32 was used to calculate genotypic and allelic frequencies of each variation and Hardy Weinberg Equilibrium (HWE) was also tested by calculating Chi^2 (Tables 3, 4 and 5). Out of seven SNPs, p.A30V, p.Y205L, p.T312M and p.A313T were obeying HWE.

2.4. CAPS

p.A313T was selected for illustrating higher values towards fixed genotypic and allelic frequencies ($P = 0.17016 > 0.05$). This variation was further tested at larger population level (n = 146). For this purpose, Cleaved Amplified Polymorphic Sequences (CAPS) was used to on larger data set. 146 more animals were selected from livestock farms and restriction digestion was performed. CviAII was used to digest the amplified DNA fragment carrying target variation (p.A313T) and results were read on agarose gel electrophoresis (Fig. 2). After this screening, genotypic frequencies were calculated and association analysis was performed by one way ANOVA and results found are mentioned in Tables 3, 4 and 5.

2.5. Homology modeling of CYP11b1 protein

Three dimensional protein model was predicted by using PyMOL (www.pymol.org) and Phyre2 (www.sbg.bio.ic.ac.uk/phyre/html/). *CYP11b1* protein is 503 amino acid long and has been well characterized previously (Fan and Papadopoulos 2013). The predicted model was further analyzed for location of probable functional domains, signal peptides, secondary structures and transmembrane helix (Figs. 3 to 7).

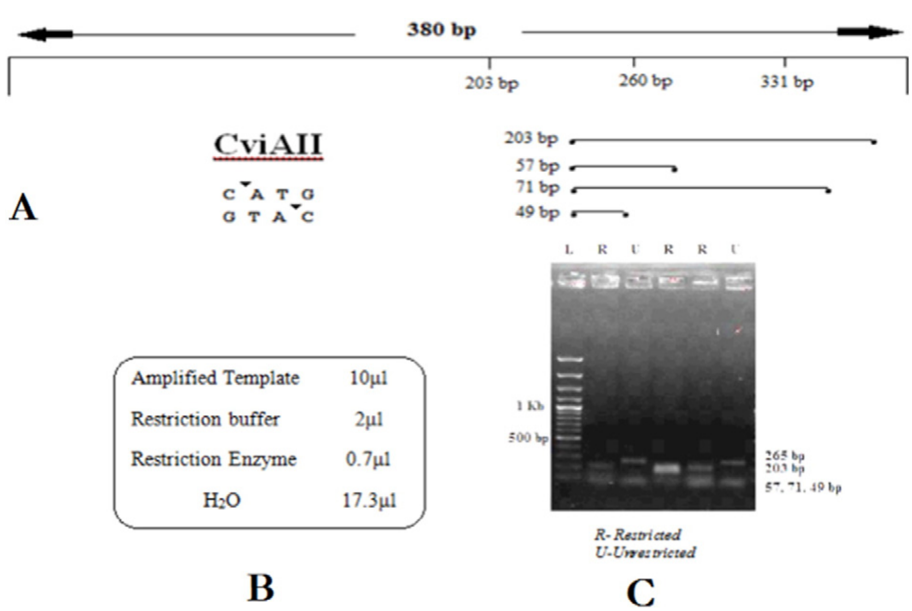

Fig 2. Cleaved Amplified Polymorphic Sites (CAPS) of p.A313T. A. Restriction mapping of variation depicting cleavage site. B. Recipe of reaction mixture used for enzymatic digestion. C. Electrophoretic pictorial of restricted DNA fragments. (Agarose gel—2.5%).

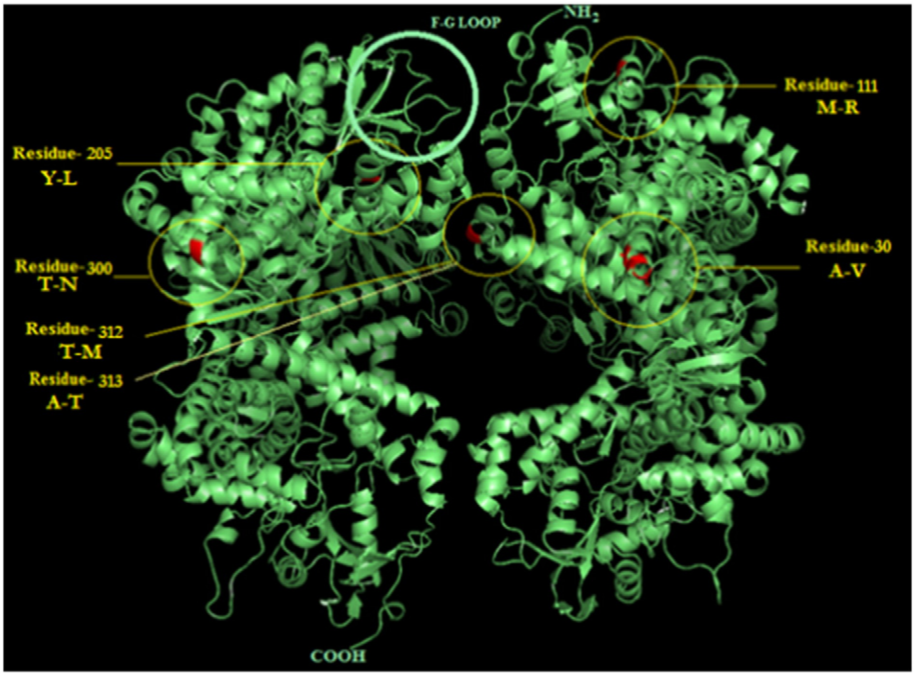

Fig 3. Structural attributes of *CYP11b1* protein. Red color is for location of polymorphic site. Residue-205 is found to be in vicinity of F–G loop, which is a major site of attachment for *CYP11b1* protein with inner mitochondrial membrane.

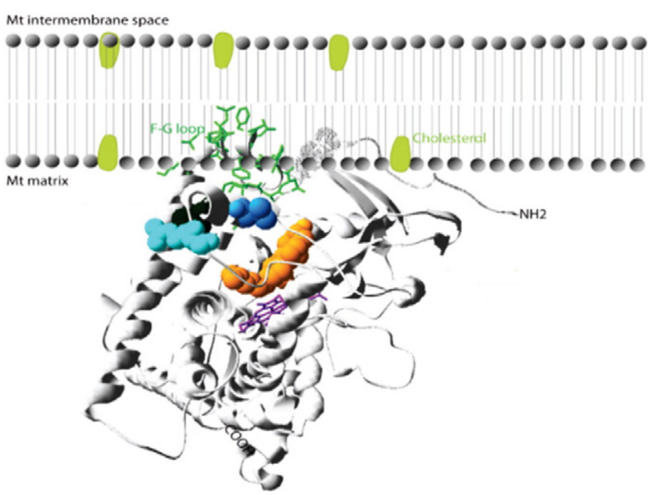

Fig. 4. F–G loop region in *CYP11b1* protein (Fan and Papadopoulos 2013).

3. Results and discussion

CYP11b1 gene is functional candidate for imparting its role in determining milk quality traits (fat in particular). This is the major enzyme in steroid genesis by converting cholesterol into conrtisol. Many of previous reports have been informative for its significance in bovines (Bu¨low and Bernhardt, 2002; Mellon et al., 1995; Zhang and Miller, 1996; Kirita et al., 1990; Okamoto et al., 1995; Sun et al., 1995; Bu¨low et al., 1996; Boon et al., 1997; Muller, 1998). In present study, *CYP11b1* was sequenced in riverine buffalo breed of Pakistan to find genomic causes of phenotypic variations identified in dairy buffalo population especially for milk content. Seven novel variants were identified by comparing the sequences of buffalo groups with higher and lower fat content (8% threshold value). Only one out of these variations was intronic. Remaining all were exonic and non-synonymous causing amino acid substitutions (Table 2). These variations were then analyzed statistically for genotypic and allelic frequency and HWE test (Tables 3 and 4). HWE testing illustrated that only three SNPs were obeying Hardy Weinberg Equilibrium and were candidate for association analysis. Out of these variations, p.A313T was further screened on larger data set of buffalo population ($n = 146$) by restriction digestion using CviAII (Fig. 2) and statistically analyzed. Single marker association was performed by one way ANOVA (Mean $\pm$ SE). AB genotype of p.A313T was found to be strongly associated with higher milk fat %age (8.9) (Table 5). Similar values for frequencies and association were calculated as given in Tables 3, 4 and 5. Finally 3D protein model for *CYP11b1* protein was predicted by Phyre2 and PyMol. Only one variation p.Y205L (residue-205) was found in vicinity of functionally important region F–G loop (Figs. 3 and 4). As this was a hydrophobic amino acid substitution (Table 6) so this might have a role in enhanced attachment of *CYP11b1* protein to inner mitochondrial membrane to start steroidogenesis. This variation is not found strongly associated on data set so far but need validation on larger buffalo population (Table 5). Secondary structural attributes of this protein have been mentioned in Fig. 5. There was one transmembrane helix as well found between residues-236 to 251 (Fig. 6). This was also not found to carry identified variations. p. A30V was found in the signal peptide near cleavage site but substitution of alanine to valine was not functionally significant or meaningful (Fig. 7).

Table 6
Polymorphism present in the vicinity of F–G loop.

Substitution (residue-205)	Group	Effect
Y–L	NonPolar	Hydrophobic

Results of present study are disparate from preceding research conducted on German Holstein cattle by Kaupe et al. (2014). They identified A30V in signal peptide of *CYP11b1* protein. This was found strongly associated with milk fat content. Same variation was also studied by Boleckova et al. (2012) in Czech Fleckvieh cattle population and was found significantly associated with milk yield and fat content. But in our study p.A30V was although found in Hardy Weinberg Equilibrium ($P > 0.05$) but could not be associated. Validation and confirmation on

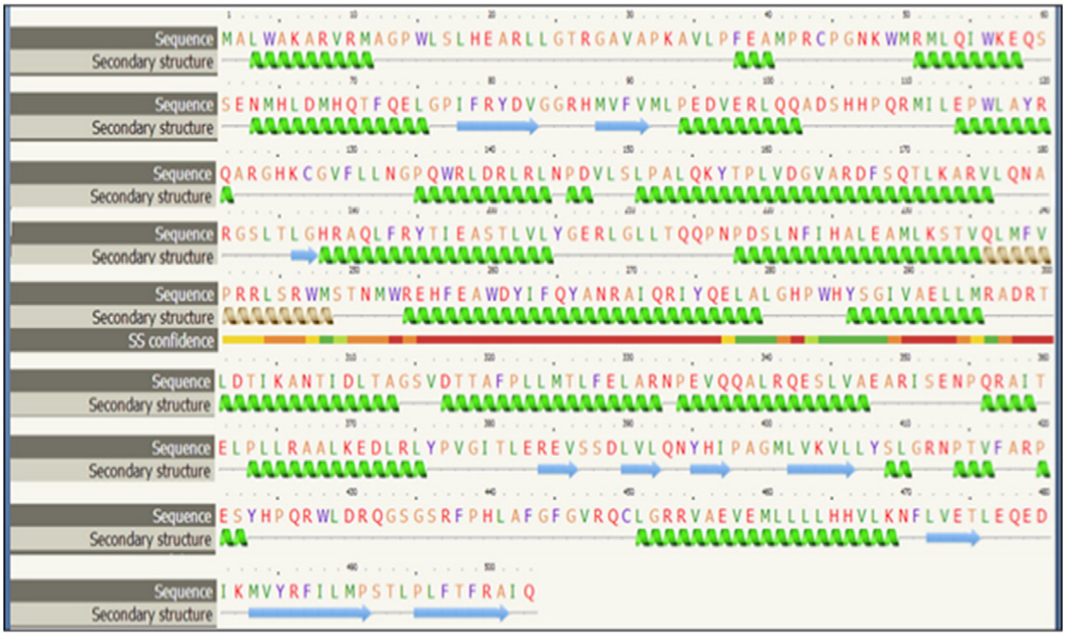

Fig. 5. Secondary structure information of *CYP11b1* protein. Green color illustrates alpha helices. Blue arrows are indicators of beta sheets.

comparatively larger set of population might be needed to study actual genetic pictorial of this locus in buffaloes. From literature, it was found that *CYP11b1* is orthologous of PPKAR1A s (Kirita et al., 1990; Bu¨low and Bernhardt, 2002). Gene duplication events occurred during vertebrate evolution resulted in only one functional *CYP11B1* gene in bovines (Kirita et al., 1990; Okamoto et al., 1995; Sun. et al., 1995; Bu¨low et al., 1996; Boon et al., 1997; Muller, 1998). Polymorphism p.Y205L was found in the vicinity of the region involved in steroidogenesis and was a hydrophobic amino acid substitution. As this region called F–G Loop is hydrophobic so this substitution might have a role in better attachment of the protein to inner mitochondrial surface and result in better protein function. Bionaz and Loor (2008) also reported effect of polymorphisms on the function of protein. Papadopoulos and Miller reported their work in 2012 on role of mitochondria in steroidogenesis. *CYP11b1* variation found significantly associated in Nili-Ravi buffalo population was p.A313T (AB genotype—8.9% fat %age). This is novel locus and has not been reported previously. Nili–Ravi buffaloes have been found heterologous for this particular locus. These results are extenuating the cross breeding of Nili and Ravi buffalo breeds almost a century back. The concentration of valuable alleles resulted in better dairy capabilities of Nili–Ravi buffalo due to phenomenon of heterosis and additive gene action. These loci provide proficient potentials in selection of genetically superior dairy buffaloes by strongly associated selection signatures. Exploration of genetic potential of animals for traits of economic importance will lead towards the identification of better genotypes at population level that can be used in future breeding program for selection of animals with superior genetic makeup.

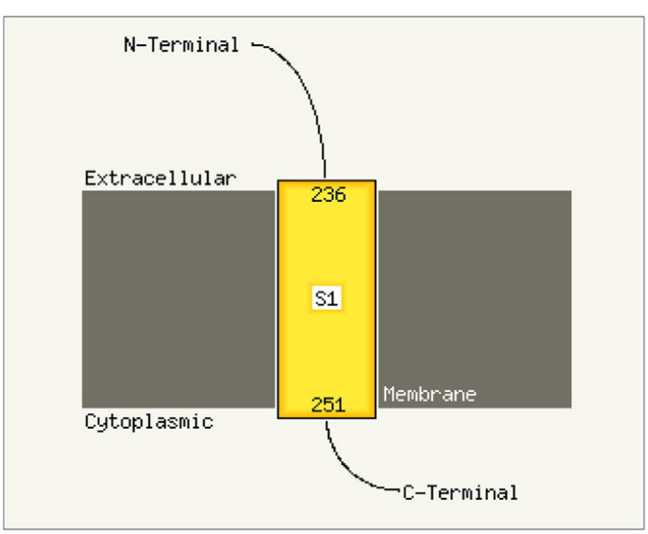

Fig. 6. Transmembrane helix of *CYP11b1*protein. This part gets attached with inner mitochondrial membrane.

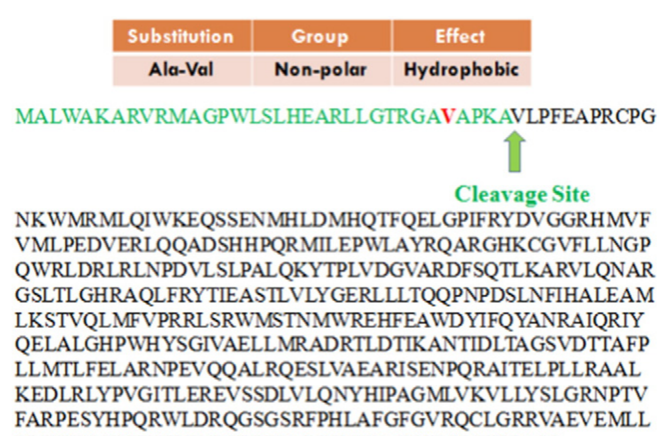

Fig. 7. Signal peptide of *CYP11b1* protein. This is 34 residues long and mutation has been identified near cleavage site.

References

Afzal, M., 2010. Re-designing smallholder dairy production in Pakistan. Pak. Vet. J. 30 (3), 187–190.

Bilal, M.Q., Sajid, M.S., 2005. Meeting milk demand (the only way is to modernize dairy farming). The Nation (May. 29: 2005).

Bionaz, M., Loor, J.J., 2008. Gene networks driving bovine milk fat synthesis during the lactation cycle. BMC Genomics 9 (1), 366.

Boleckova, J., Matejickova, J., Stipkova, M., Kyselova, J., Barton, L., 2012. The association of five polymorphisms with milk production traits in Czech Fleckvieh cattle. Czech J Anim Sci. 57 (2), 45–53.

Boon, W.C., Roche, P.J., Butkus, A., McDougall, J.G., Jeyaseelan, K., Coghlan, J.P., 1997. Functional and expression analysis of ovine steroid II beta-hydroxylase (cytochrome P450 II eta). Endocr. Res. 23, 325–347.

Brettes, J.P., Mathelin, C., 2008. Effet dual des androgenssur la glande mammaire. Bull. Cancer 95, 495–502 (Paris).

Bu¨low, E.H., Bernhardt, R., 2002. Analyses of the *CYP11B1* gene family in the guinea pig suggest the existence of a primordial *CYP11B1* gene with aldosterone synthase activity. Eur. J. Biol. Chem. 269, 3838–3846.

Bu¨low, E.H., Mo¨bius, K., Ba¨hr, V., Bernhardt, R., 1996. Molecular cloning and functional expression of the cytochrome P450 11B hydroxylase of the guinea pig. Biochem. Biophys. Res. Commun. 221, 304–312.

Fan, J., Papadopoulos, V., 2013. Evolutionary origin of the mitochondrial cholesterol transport machinery reveals a universal mechanism of steroid hormone biosynthesis in animals. PLoS. ONE. 8, 10.

Kaupe, B., Kollers, S., Fries, R., Erhardt, G., 2004. Mapping of CYP11B1 and a putative CYHR1 paralogous gene to bovine chromosome 14 by FISH. Anim. Genet. 35, 6.

Kirita, S., Hashimoto, T., Kitajima, M., Honda, S., Morohashi, K., Omura, T., 1990. Structural analysis of multiple bovine P-450 (11 beta) genes and their promoter activities. J. Biochem. 108, 1030–1041 (Tokyo).

Maryam, J., Babar, M.E., Nadeem, A., Hussain, T., 2012. Genetic variants in interferon gamma (IFN-γ) gene are associated with resistance against ticks in *Bos taurus* and *Bos indicus*. Mol. Biol. Rep. 39 (4), 4565–4570.

Mellon, S.H., Bair, S.R., Morris, H., 1995. P450cllB3 mRNA, transcribed from a third p450cll gene, is expressed in a tissue specific, developmentally, and hormonally regulated fashion in the rodent adrenal and encodes a protein with both 11-hydroxylase and 18-hydroxylase activities. J. Biochem. 270, 1643–1649 (Tokyo).

Muller, J., 1998. Regulation of aldosterone biosynthesis: the end of the road? Clin. Exp. Pharmacol. Physiol. Suppl. 25, S79–S85.

Okamoto, M., Nonaka, Y., Ohta, M., Takernori, H., Halder, S.K., Wang, Z.N., Sun, T., Hatano, O., Takakusu, A., Murakami, T., 1995. Cytochrome P450 (11 beta): structure-function relationship of the enzyme and its involvement in blood pressure regulation. J. Steroid Biochem. Mol. Biol. 53, 89–94.

Payne, A.H., Hales, D.B., 2004. Overview of steroidogenic enzymes in the pathway from cholesterol to active steroid hormones. Endocr. Rev. 25, 947–970.

Sun, T., Zhao, Y., Nonaka, Y., Okamoto, M., 1995. Cloning and expression of cytochrome p450 (11 β) of porcine adrenal cortex. J. Steroid Biochem. Mol. Biol. 52, 227–232.

Zhang, G., Miller, W., 1996. The human genome contains only two *CYP11B1* (P450c11) genes. J. Clin. Endocrinol. Metab. 81, 3254–3256.