Research Article **Biotechnology**

International Journal of Pharma and Bio Sciences

ISSN 0975-6299

MOLECULAR DIFFERENTIATION OF MITOCHONDRIAL GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE AMONG DIFFERENT BREEDS OF *BUBALUS BUBALIS*

ANIL KUMAR[1], VINOD CHHOKAR[*1], RAVINDER KUMAR[1], VIKAS BENIWAL[1], S. N. KALA[2], ANAMIKA MISHRA[3], ASHWIN ASHOK RAUT[3] AND P. K. NAIK[4]

[1]*Department of Bio & Nano Technology, Guru Jambheshwar University of Science & Technology, Hisar-125001, Haryana, India.*
[2]*Central Institute for Research on Buffalo, Hisar-125001, Haryana, India.*
[3]*High Security Animal Disease Laboratory, Bhopal-462021, Madhya Pradesh, India.*
[4]*Department of Bioinformatics & Biotechnology, Jaypee University of Information Technology, Waknaghat, Solan-173215, Himachal Pradesh, India.*

ABSTRACT

Mitochondrial Glycerol-3-phosphate acyltransferase (GPAM; EC 2.3.1.15), catalyzes the first step of glycerolipid biosynthesis. It plays a key role in the regulation of cellular triacylglycerol and phospholipid levels. The GPAM esterification step constitutes the committed reaction in the de novo synthesis of triacylglycerol (TAG) and all the acyl-glycerol phospholipids. In the present investigation GPAM was partially amplified (around 652 bp) by designing gene specific primers and confirmed by sequencing the amplicon and its comparison with the GPAM gene of bovine. Comparative study of GPAM among different breeds of buffaloes reveals different level of mutations with respect to its gene sequence 2.7-5.6% and protein sequence 3.4% to 13%. Similarly the protein structures modeled from their sequences were compared by structural superposition that shows variations (Root Mean Square Deviation) from 0.26 to 1.025. Molecular differentiation among different breeds of buffaloes reveals that the reference and *Bhadawari* are very close to each other, *Toda* and *Surti* are similarly close to each other and *Pandharpuri* is related to *Murrah*. Especially the variations are more at the binding site of this protein that may be the cause that different breeds have different percentage of milk fat. Further study is underway to detect polymorphism and associate them with milk fat related traits in buffalo.

KEYWORDS: **GPAM**, ***Bubalus bubalis***, ***Murrah***, ***Surti***, ***Pandharpuri***, ***Bhadawari***, ***Toda***, **molecular differentiation**

VINOD CHHOKAR

Department of Bio & Nano Technology, Guru Jambheshwar University of Science & Technology, Hisar-125001, Haryana, India.

INTRODUCTION

The global population of water buffalo has increased 98% in the last decade, from 88 million in 1961 to 188 million in 2009. With this large population of buffalo which is very well adapted to the tropical environment and has a potential to meet the protein requirement of our growing human population, it appears quite appropriate to call it "Black gold". Buffalo contribute 57% of the total milk production in India. Buffalo farming is increasing in India due to the growing market demand for buffalo milk. The quality of milk based on fat composition varies significantly among different breeds of buffalo. For example the breed, *Murrah* consists of average 6.7% milk fat while *Bhadawari* contains average13% milk fat[1]. About ten breeds of buffalo: *Murrah*, *Pandharpuri*, *Bhadawari*, *Surti*, *Toda*, *Nili-Ravi, Jaffarabadi, Mehsana, Marathwada, Nagpuri* are reported in India. Out of them *Murrah* gives maximum amount of milk (~18-20ltr), however, the quality of the milk is not good because of low fat content. This difference in milk fat percentage across the different breeds clearly indicates that there is an involvement of a genetic factor. It also implies that over the evolutionary period the genes affecting milk fat percentage have segregated and fixed differently in the various breeds of buffaloes. Since within a species the basic physiological processes are more or less the same, this difference in production potential can be attributed only small variation. These small differences which probably the differences in the allelic distribution of the genes involved in the milk fat synthesis are resulting in remarkable differences of production potential. The knowledge of these small genetic differences would be invaluable as they can be used to enhance the rate of genetic improvement.[1]

With the development of refined analytical tools in the genome analysis, the identification of causal genes and their sequence variation underlying complex traits has become feasible in humans, animals, and plant[2]. Milk fat percentage is a quantitative trait that is determined by collective effect of multiple gene and environmental factor. Based on present knowledge of physiology of milk fat synthesis the major candidate gene that influences milk fat percentage is mitochondrial glycerol-3-phosphate acyltransferase Glycerol-3-phosphate acyltransferase (GPAM) is one of the most important enzyme that regulates the milk fat synthesis. It catalyzes the first step of glycerolipid biosynthesis. It plays a key role in the regulation of cellular triacylglycerol and phospholipid levels. There are two isoforms of glycerol-3-phosphate acyltransferase in mammals, a mitochondrial and a cytosolic form. The mitochondrial form (GPAM) prefers saturated fatty acyl-CoA as a substrate, whereas the cytosolic enzyme uses both saturated and unsaturated fatty acyl-CoA[3]. The synthesis of triacylglycerol and the glycerophospholipids begins with the acylation of glycerol-3-phosphate by glycerol-3-phosphate acyltransferase (GPAT) to form LPA , a rate-limiting step catalyzed by one of several isoforms of glycerol-3-phosphate acyltransferase (GPAT, EC 2.3.1.15)[4]. In this study attempt has been made to study the molecular differences in GPAM among different breeds of buffalo that have different percentage of milk fat. Furthermore, the impact of variations in gene sequences if any detected among the different breeds onto protein structure and its binding site was also assessed.

MATERIALS AND METHODS

(i) Samples collection and DNA isolation

Five different breeds of buffaloes were used i.e. *Murrah*, *Pandharpuri*, *Bhadawari*, *Surti*, and *Toda* that were reared in different agro climatic regions of India. *Murrah* is the major buffalo breed from northern India whereas *Pandharpuri* is from western India. *Bhadawari*

is the major breed of Agra and Etawa (Utter Pradesh) Bhind and Morena (Madhya Pradesh). *Surti* is major breed from north Gujarat and Rajasthan. *Toda* is major breed of Neelgiri hills of Madras (Tamilnadu).[1] High milk fat yields and breed-specific traits were the main selection among these breeds—*Bhadawari* composed of high milk fat (13%) followed by *Toda* (8.2%), *Surti* (8.10%), *Pandharpuri* (7.0%) and *Murrah* (6.7%). Only healthy animals with the characteristic traits of each breed were selected. Approximately 10 ml venous blood was collected from each breed with 0.5 ml of 2.7% EDTA as anticoagulant. The samples were brought to the laboratory in a double walled ice box containing ice with cool pack and stored at-20^0C till the isolation of DNA. Genomic DNA was isolated from buffy coat of venous blood sample of buffalo using phenol chloroform extraction method[5]. The quality of DNA was checked on spectrophotometer quality DNA was used after an appropriate dilution for further work.

(ii) Primer design and amplification of GPAM

GPAM from bovine has been well characterized but not in buffalo. It consists of 20 numbers of exon with 6196 total base pairs; exon 20 is the largest one consisting of 3739 base pairs. The amino acid sequence of GPAM spanning within 10 Å of active site of the enzyme was extracted and its corresponding gene sequence was used for the primer designing. This region composed of approximately 652 bp of exon 20 and covered 268 bp to 920 bp as mentioned in Figure 1. A set of forward primer consisting of 21bp 5’TGATCTGTGATCGCTGGGCAC3’ and reverse primer with 21 bp 5’GCTCCCAAACACTTGCCTGTC 3’ were designed by using the *Bos taurus* sequence available on the NCBI. For amplification of this partial sequence of GPAM, 50 µl of PCR reaction was prepared by adding 10pM of each primer, 0.2µM of each dNTPs, 1.5mM $MgCl_2$, 10xPCR assay buffer, 130ng DNA template and 1 unit *Taq.* DNA polymerase. The amplification was carried by using a programmable thermal controller (PTC-100, MJ Research) with the following conditions: initial denaturation of 5 minute at 94^0C followed by 30 cycles of denaturation at 94^0C for 45 second, annealing at 58.6^0C for 45 second and extension at72^0C for 45 second and finally the extension of 10 min. at 72 ^{0}C. For sequencing, the PCR products were run in 1.5%agarose gel and product band was eluted using QIAquick gel extraction kit. The amplified products were sequenced by Sanger method by outsourcing (Xclaris lab Ahmedabad (India).

(iii) Comparison of DNA and protein sequences

The gene sequences obtained based on forward and reverse primers from different breeds of buffaloes were assembled utilizing overlapping sequences (consisting of 652 bp) and were compared. These sequences were aligned using multiple alignment programs, Clustal W (http://www.ebi.ac.uk/Tools/msa/clustalw2/) with DNA weight matrix and multiple parameters like gap opening 10.0, gap extension 0.20, transition weight 0.50. A phylogenetic tree based on similarity coefficients generated by neighborhood-joining method was performed. Furthermore, all DNA sequences obtained were translated to protein sequence using EMBL transseq (http://www.ebi.ac.uk/Tools/st/). These protein sequences were further aligned using multiple alignment programs, Clustal W (http://www.ebi.ac.uk/Tools/msa/clustalw2/) with protein weight matrix and default parameters for gap opening and extension. A phylogenetic tree based on similarity coefficients was also generated.

(iv) Modeling of protein structure

The protein sequences were BLAST with PDB data base that would give only one similar protein with sequence similarity of 62% (PDB ID: 1Q14). This homolog protein structure was used as template for homology model building of GPAM using Prime accessible through the

This article can be downloaded from www.ijpbs.net

Maestro interface (Schrodinger, Inc.). All water molecules were removed for the template. During the homology model building Prime kept the backbone rigid for the cases in which the backbone did not need to be reconstructed due to gaps in the alignment. The model was screened for unfavorable steric contacts and remodeled using a rotamer library database of Prime. Explicit hydrogen was added to the protein and the protein model was subjected to energy minimization using the Macromodel (Prime version 1.5)[6] force-field OPLS-2005. Energy minimization and relaxation of the loop regions were performed using 300 iterations in a simple minimization method. The steepest descent energy minimization was carried out until the energy showed stability in the sequential repetition. Model evaluation was performed in PROCHECK v3.4.4 producing plots which were analyzed for the overall and residue-by-residue geometry[7]. Ramachandran plot[8]; provided by the program PROCHECK assured very good confidence for the predicted protein. Nevertheless, PROCHECK assured the reliability of the structure and the protein was subjected to VERIFY3[9]Davailable from NIH MBI Laboratory Servers.

(v) Prediction and characterization of binding site

For each protein structure that had been modelled, the binding sites were defined as a set of site points that were predicted using 'SiteMap' (Schrödinger Inc., version 2.4). A SiteMap calculation has three stages. In the first, relevant site points are selected based on geometric and energetic properties and the points are grouped in sets to define the binding sites. Next, hydrophobic, hydrophilic, and other key properties are computed at grid points and contour maps are prepared. Finally, binding site properties such as site score, size, volume, enclosure, exclosure, contact, hydrophobic, hydrophilic and ratio of hydrogen donor to acceptor are computed. Finally, each binding site is scored by an expression that uses just three terms: (1) the square-root of the number of site points, capped at 100 site points to avoid overly rewarding large sites; (2) the enclosure score, and (3) the hydrophilic score. The 'Site Score' function is represented by the following equation.

$$\text{Site Score} = 0.0733\, n^{1/2} + 0.6688\, e - 0.20\, p \quad (1)$$

Where n is the number of site points (up to 100), e is the enclosure score, and p is the hydrophilic score, capped at 1.0 (the average for the submicromolar sites) to limit the impact of hydrophilicity in charged and highly polar sites.

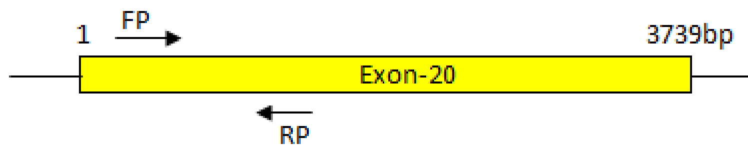
RESULTS AND DISCUSSION

Genomic DNA from blood samples was isolated as per the method of Sambrook and Russel[5]. The quality of genomic DNA was checked in 0.8% submarine agarose gel electrophoresis. Most of the DNA samples were of good quality and exhibited single band, where as few samples of poor quality DNA showed a smear throughout the lanes. Smearing of the DNA was found either due to the shearing of DNA or presence of protein in the sample. Those samples were re-extracted by phenol-chloroform method

(i) Amplification of GPAM gene

GPAM gene from different breeds of buffaloes were partially amplified (~ 652 bp) using the primer specifically designed for GPAM (Figure.1a). A clear, high reproducible and distinct band of size 652 bp was obtained (Figure.1b). The amplified products were confirmed by nucleotide sequencing using the same set of primers and by comparing with the reference gene of GPAM from bovine. The amplified fragment of GPAM corresponded to part of Exon 20.

(a)



Primer	Sequence	Position	T_m (^{0}c)	Product size
	5' TGATCTGTGATCGCTGGGCAC 3'	268	71.2	652
P1	5' GCTCCCAAACACTTGCCTGTC 3'	920		

(b)

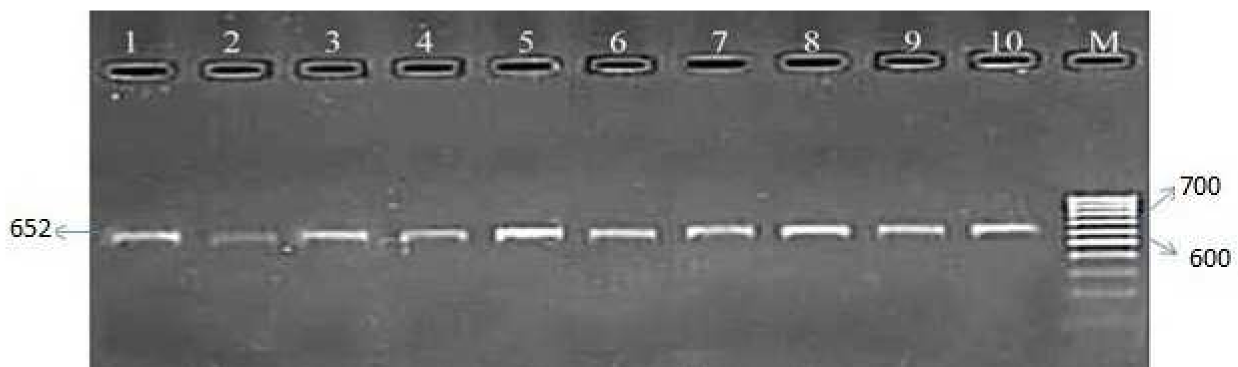


Figure1.
Primer and partially amplified GPAM from different breeds of buffaloes. A gene specific primer was designed based on Exon 20 sequence from bovine with a product size of 652 bp out of total length of 3739 bp (a). A representative gel picture showing the amplified product of GPAM from different breeds of buffaloes (b): 1 & 2, Murrah; 3 & 4, Pandharpuri; 5 & 6, Bhadawari; 7& 8 Surti; 9 & 10 Toda.

(ii) Sequence comparison of GPAM

A relatively small variation was detected among the GPAM gene sequences obtained from different breeds of buffaloes. All the gene sequences were compared using multiple sequence alignment programs and level of variation in gene sequences is represented in Figure. 2a. Genomic variation among various breeds' ranges from 2.7-5.6% in comparison to GPAM sequence obtained from bovine. The breed, *Bhadawari* had the minimum percentage of variation (2.7%) followed by *Murrah* (3.3%), *Toda* (4.6%), *Surti* (4.7%) and *Pandharpuri* (5.6%).Furthermore, the sequences were used to generate a dendrogram (Figure 2b). It was seen that the reference and *Bhadawari* are very close to each other, *Toda* and *Surti* are similarly close to each other and *Pandharpuri* is related to *Murrah*.

(a) Multiple sequence alignment of GPAM

(b) Relationship between GPAM among different breeds

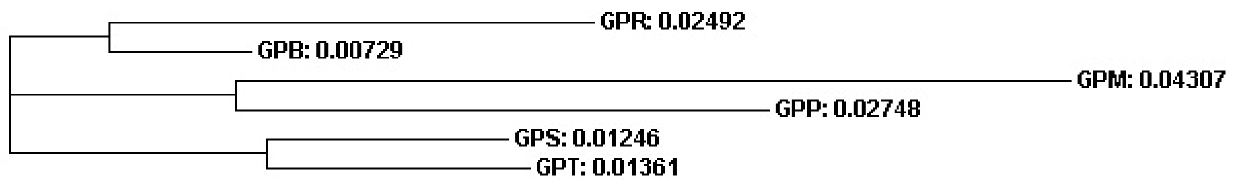


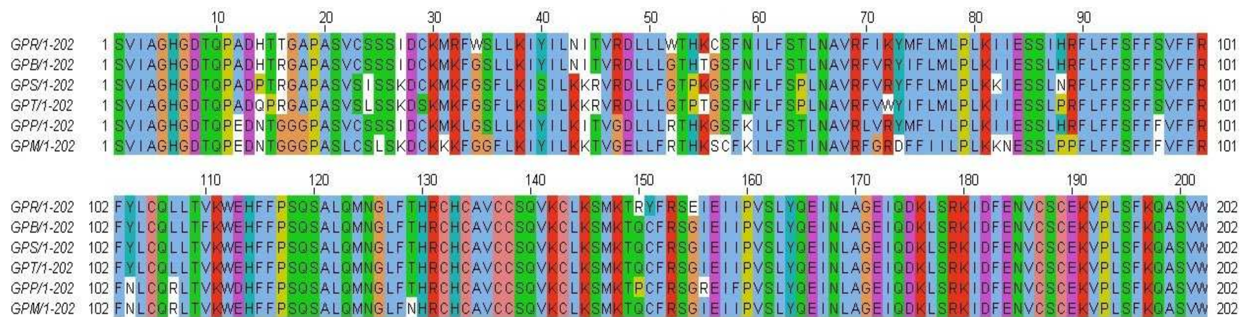
Figure 2.

Comparison of GPAM gene sequences. (a) The GPAM sequences obtained from different breeds of buffaloes were aligned using multiple sequence alignment programs, the changed nucleotide bases were represented in white box. Only few mutations were detected between different breeds. (b) The dendrogram representing the similarity between GPAM gene sequences obtained from different breeds.

Based on the comparison of protein sequences of GPAM, a relatively small variation was detected among the different breeds of buffaloes (Figure. 3a). Overall the rate of variation ranges in between 3.4% to 13%, with *Murrah* having the highest variation (13%),

followed by *Toda* (9.3%), *Pandharpuri* (7.8%), *Surti* (7.3%) and *Bhadawari* (3.4%) in comparison to the reference sequence of bovine. The dendrogram (Figure. 3b) obtained from pairwise comparison of protein sequence also reveals similar findings as mentioned above. Both *Bhadawari* and reference are very close to each other, *Toda* and *Surti* are similarly close to each other and *Pandharpuri* is related to *Murrah*. The variation in gene as well as protein sequences of GPAM might be associated with variation in milk fat content.

(a) Multiple sequence alignment of GPAM



(b) Relationship between GPAM among different breeds

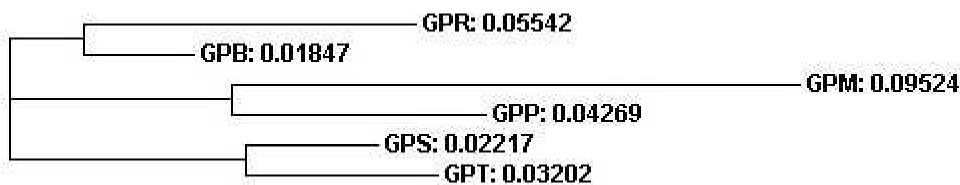


Figure 3

Comparison of GPAM protein sequences. (a) The GPAM sequences obtained from different breeds of buffaloes were aligned using multiple sequence alignment program, the changed amino acids were represented in white box. Only few mutations were detected between different breeds. (b) The dendrogram representing the similarity between GPAM protein sequences obtained from different breeds.

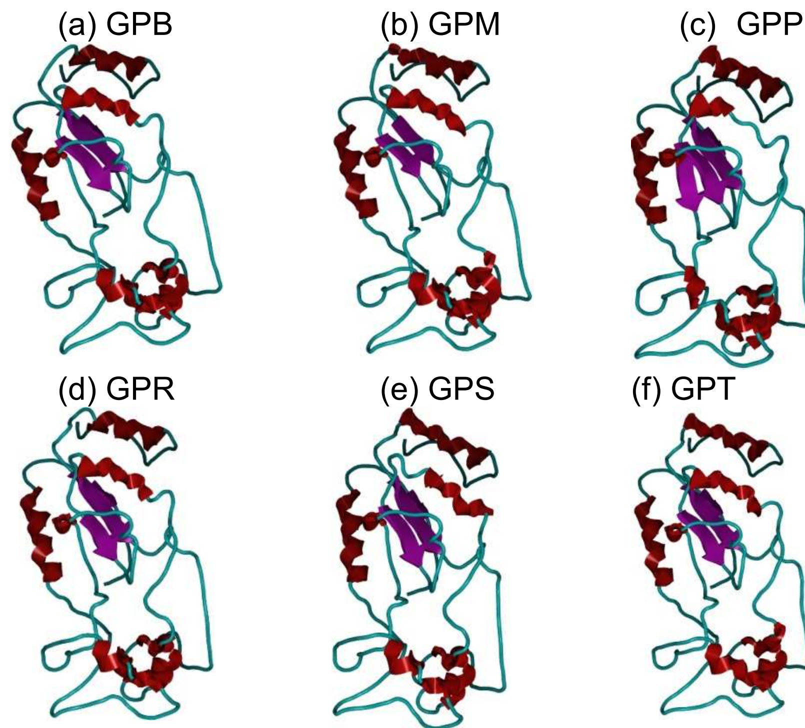
(iii) Modeling of GPAM protein structure and comparison among different breeds

The atomic coordinates of GPAM protein structure for the organism *Bubalus bubalis* was not available in Protein Data Bank which necessitated for developing a protein model. The final model, which we modeled and took for further analysis, consisted of 202 amino acid residues. We used both PROCHECK and VERIFY3D softwares to check the quality of the modeled protein. Ramachandran plot obtained from the program PROCHECK, which checks the stereochemical quality of a protein structures, producing a number of postscript plots analyzing its overall and residue-by-residue geometry, assured the reliability of the modeled protein with 87% residues in most allowed region and 7.7% in additional allowed region[10]. There were only 1.2% residues in disallowed region and 4.1% in generously allowed region. The assessment with VERIFY3D, which derived a “3D-1D” profile based on the local environment of each residue, described by the statistical preferences for the area of the residue that was buried, the fraction of side-chain area that

This article can be downloaded from www.ijpbs.net

was covered by polar atoms(oxygen and nitrogen) and the local secondary structure, also substantiated the reliability of the three dimensional structure. The residues that deviated from the standard conformational angles of Ramachandran plot were the members of N terminal domain of the protein. This was an ignorable condition since the N-terminal end was not critical in our study. The distance of these residues to the active site residues also were found to be more than 10 Å, which suggested that those residues would interfere little with the binding of ligands in the active site region of GPAM. The structural comparison of GPMA from different breeds revealed small variation among each other (Figure. 4a). The pairwise comparison of protein structure (expressed as root mean square deviation, RMSD) varies from 0.26-1.025,the minimum RMSD was between *Murrah* and reference (0.26), where as the maximum value was found between *Pandharpuri* and *Toda*(1.025)(Figure.4b). Furthermore, the dendrogram generated based on pairwise matrix of RMSD corroborate similar findings, *Bhadawari* and reference are closely related, *Toda* and *Surti* are similarly close to each other and *Pandharpuri* is related to *Murrah* (Figure. 4c). Various mutations detected in the gene and protein sequence of GPAM among different breeds of buffaloes have different level of impact in the binding site (Table 1). We have predicted the binding sites using SiteMap algorithm (Schrodinger software package). However, we considered only the best binding site (based on Site Score) for the comparative study. The characteristic features of binding site analyzed using SiteMap is included in Table 1. The significant variation in binding site among different breeds of buffaloes may be associated with variation in milk content.

(a) The carton structure of GPAM protein modeled using homology based technique from different breeds. The protein structures are more or less similar to each other.



This article can be downloaded from www.ijpbs.net

(b) Pairwise RMSD between protein structures

	GPR	GPM	GPP	GPB	GPS	GPT
GPR	0					
GPM	0.26	0				
GPP	0.855	0.952	0			
GPB	0.638	0.819	0.775	0		
GPS	0.518	0.660	1.003	0.834	0	
GPT	0.55	0.683	1.025	0.883	0.392	0

(c) Dendrogram showing relatedness between protein structures

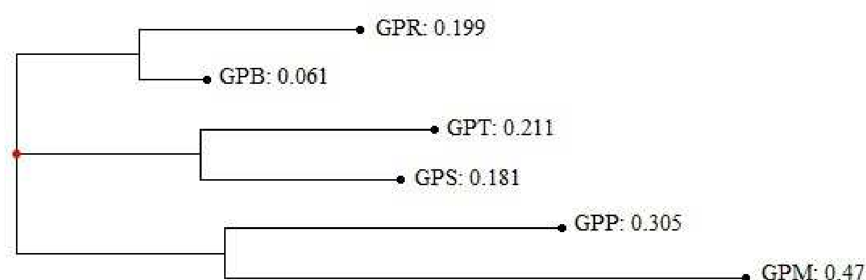


Figure 4.
The modeled protein structures of GPAM and its comparison. (a) The carton structure of GPAM protein modeled using homology based technique from different breeds. The protein structures are more or less similar to each other. (b) The pairwise comparison of modeled GPAM protein structure from different breeds. The value represents root mean square deviation (RMSD). (c) A dendrogram representing the similarity between protein structures of GPAM from different breeds. The tree was constructed using pairwise RMSD matrix.

Table1
The binding site of GPAM (predicted using Schrodinger software package) and its physico-chemical properties

Proteins	Site Score	Size	Volume	Exposure	Enclosure	Contact	Hydrophobic	Hydrophilic	Donor/ Acceptor
GPR	1.054	275	700.7	0.506	0.743	0.981	0.987	0.883	0.740
GPM	1.087	334	833.5	0.476	0.787	1.004	1.301	0.863	0.465
GPP	1.125	157	331.3	0.537	0.793	1.040	1.372	0.695	0.772
GPB	1.085	399	881.2	0.580	0.698	0.876	1.521	0.574	0.825
GPS	1.059	283	760.1	0.557	0.728	0.925	1.169	0.807	0.582

This article can be downloaded from www.ijpbs.net

REFERENCES

1. Kumar A, Kumar R, Beniwal V, Kala S N, Mishra A, Raut A A, Naik P K and Chhokar V, Molecular differentiation of Peroxysome proliferator activated receptor coactivator-1 among different breeds of Bubalus bubalis, Bioinformation, 8(13):600 – 606,(2012).
2. Glazier, A.M., Nadeau, Joseph H., and Aitman Timothy J, Finding genes that underlie complex traits. Science 298: 2345–2349, (2002)
3. Roy R,. Gautier M, Hayes H, Laurent P, Zaragoza P, Eggen A and Rodellara C, Assignment1 of mitochondrial glycerol-3-phosphate acyltransferase (GPAM) gene to bovine chromosome 26 (26q22) by in situ hybridization and confirmationby somatic cell hybrid mapping, Cytogenet Genome Res, 97:276–277, (2002).
4. Gonzalez-Baro MR, Lewin TM and Coleman RA, Regulation of triglyceride metabolism: II, function of mitochondrial GPAT1 in the regulation of triacylglycerol biosynthesis and insulin action. Am J PhysiolGastrointest, Liver Physiol, 292:G1195–G1199, (2007).
5. Sambrook J and Russel DW, Ed. Molecular cloning-A laboratory manual, 3rd Edn, Cold Spring Harbor Laboratory Press, New York, (2001).
6. Prime version 1.5, Macromodel version 9.1, Schrodinger, LLC,New York, NY, (2005).
7. Laskowski RA, Mac Arthur MW, Moss DS and Thornton JM, PROCHECK: a program to check the stereochemical quality ofprotein structures, J Appl Crystallogr, 26:283–291, (1993).
8. Ramachandran GN, Ramakrishnan C and Sasisekharan V, Stereochemistry of polypeptide chain conFigureurations, J Mol Biol, 7:95–99, (1963).
9. Eisenberg D, Luthy R and Bowie JU, VERIFY3D: Assessment of protein models with three-dimensional profiles, Methods Enzymol, 277:396–404, (1997).
10 Shankaracharya., Srivastava, Bhavana and Vidyarthi Ambarish S: Structure modeling of vrk1 protein and its molecular docking study with ribavirin analogs. International Journal of Pharma and Bio Sciences.1 (3).(2010).

.

This article can be downloaded from www.ijpbs.net