Life Science Informatics Publications

Research Journal of Life Sciences, Bioinformatics, Pharmaceutical and Chemical Sciences

Journal Home page http://www.rjlbpcs.com/

RJLBPCS

ISSN 2454-6348

Original Research Article **DOI: 10.26479/2018.0406.37**

IMPACT OF DELETERIOUS SINGLE NUCLEOTIDE POLYMORPHISMS ON OVINE G6PD GENE USING COMPUTATIONAL ANALYSIS APPROACH

George P. Laliotis[1]*, Iosif A. Bizelis[2], Emmanuel Rogdakis[2]

1.Research Institute of Animal Science, Hellenic Agricultural Organization "DEMETER, Paralimni Gianitsa, 58100, Pella.

2.Laboratory of Animal Husbandry, Department of Animal Science and Aquaculture, Agricultural University of Athens, 75 Iera Odos, 11855, Athens, Greece.

ABSTRACT: Ovine Glucose 6-Phosphate Dehydrogenase (G6PD) catalyzing the first committed reaction in the pentose phosphate shunt, plays a crucial role in ruminant's lipogenesis as it contributes with the necessary reductive power by terms of NADPH in the synthesis of fatty acids. Although the respective gene has been cloned and a diminish of the its enzyme activity have been previously reported, no association study with potential mutations has been conducted. Thus, the possible functional SNPs is advisable to be sorted out, before the design of a larger population study. Using the existing databases and computational analysis, herein, 305 single nucleotide polymorphisms (SNPs) were noted with six of them reported as non-synonymous (ns) and three of them (R136C; P396L; P489S) as damaging. Further analysis of the damaging SNPs revealed a decrease of protein stability and structural changes of protein's fold, while the Proline (P) amino acid substitutions were found to be a pivotal one due to high conservation scale. In addition, a resemblance with respective deficient or disease cases of the human counterpart were observed, rendering the observed ovine nsSNPs of high importance. Results shed light into further biochemical pathways on ruminant lipogenesis and may be used for further phylogenetic or population genetics.

KEYWORDS: G6PD, mutations, SNPs, sheep, lipogenesis.

Corresponding Author: Dr. George P. Laliotis* Ph.D.

Research Institute of Animal Science, Hellenic Agricultural Organization "DEMETER, Paralimni Gianitsa, 58100, Pella. Email Address: glaliotis@rias.gr

1.INTRODUCTION

The energy balance of an organism is considered as one of the major roles of adipose tissue, providing, by terms of fatty acids, when it is required the appropriate fuels [1]. Apart from the above, the adipose tissue in productive animals may affect the economic return of a livestock, as the excess of fat deposits not only influence negatively the quality and the grading of carcasses but also influence the health status and future performance when i.e. such excesses are observed during puberty. Unlike to non-ruminant species, the predominant anatomic site for lipogenesis in ruminants is considered the adipose tissue [2,3]. The fatty acid biosynthesis apart from carbonic substrate, it also requires great amounts of NADPH in order to reduce acetyl-CoA to fatty acids. In addition, the proper reduction power by terms of NADPH, is major supplied by the pentose phosphate shunt, with the glucose 6-phosphate dehydrogenase (G6PD) to play a vital role. The certain enzyme is responsible for delivering the 50–80% of the essential NADPH compounds for the fatty acid synthesis [2], with the rest ones to be provided by the cytosolic NADP isocitrate dehydrogenase (IDH1). The advent of novel technology in the field of molecular genetics permitted the more elucidated study of genes influencing lipogenesis. Ovine G6PD cDNA and promoter region have already been cloned and characterized, while structural changes have also been noted to the ovine G6PD protein. In addition, ovine G6PD is considered a highly conserved protein compared to the respective eukaryotic counterparts [3,4,5]. The recent sequence and characterization of the whole ovine genome [6] facilitates many further studies concerning the exploitation of genetic variability of such genes. Single nucleotide substitutions are considered as a valuable source of genetic variability [7,8]. Although many types of such substitutions are observed throughout a genome, only the non-synonymous SNPs (nsSNPs), also known as missense variants, are considered of vital importance, mainly because they lead to changes in the open reading frame of translated amino acid sequence. A great number of such substitutions have been observed to the human G6PD gene reflecting to various diseases cases (deficiency) or lethal situations [9,10]. Although previous studies have been reported a potentially diminish of ovine G6PD enzyme activity [4,11,12,13], to the best of our knowledge no association with any single nucleotide substitution has been referred. Thus, the aim of the present study is, using computational approaches, to come up with any single nucleotide substitutions that will render a detrimental effect to the function, regulation and stability of the ovine G6PD protein. Such approaches could be further used in population genetics or phylogenetic analysis in association studied of the determined substitutions with other productive traits.

2. MATERIALS AND METHODS

2.1 Ovine G6PD SNPs Search

The data concerning information of ovine G6PD was collected from Entrez Gene on National Center for Biotechnology Information (NCBI) web site. SNPs information about ovine G6PD gene and their respective protein sequences was retrieved from Ensembl database

(https://www.ensembl.org/index.html) for further in silico analysis.

2.2 Evaluation of coding SNPs by sequence homology-based methods

Three web based softwares were used in order to predict functional effects of coding SNPs identified from the database search. These algorithmic programs included SIFT-software (http://sift.jcvi.org/www/SIFT_seq_submit2.html), PolyPhen-2-Polymorphism Phenotyping v2 (http://genetics.bwh.harvard.edu/pph2/) and nsSNPAnalyzer (http://snpanalyzer.uthsc.edu/). All the softwares evaluate the coding mutations in a protein sequence and are used to predict whether an amino acid substitution may have a phenotypic effect and in what extent (tolerant, benign, damaging or deleterious). The SNPs predicted as damaging/deleterious/disease by all the implemented in-silico tools were considered as high risk nsSNPs and used for further analysis.

2.3 Protein Stability

The I-Mutant v.3.0 Suite (http://gpcr2.biocomp.unibo.it/cgi/predictors/I-Mutant3.0/I-Mutant3.0.cgi) was used to estimate any potential stability damages caused by the identified missense SNPs in the protein structure. The software predicts the reliability index (RI) of the results ranging from 0–10, where 10 being the highest reliability [14]. Simulator conditions for all the examined cases were set at 25°C and pH 7.0.

2.4 Structural effects and models' creation

Project HOPE (http://www.cmbi.ru.nl/hope/) was used to search protein 3D structures, to estimate the result of the amino acid change on protein structure and to further analyze the resulting change in the 3D (tertiary/quaternary) structure [15]. MUSTER was used to create 3D protein model [16].

2.5 Evolutionary conservation, post-translational modification sites and protein-protein interactions

Estimations regarding evolutionary conservation of amino acid positions was assessed using ConSurf (http://consurf.tau.ac.il/2016/) using protein sequence [17]. Degree of conservation of amino acid residues was estimated using 150 homologous sequences. Post translational sites were predicted as follows: a) the putative ubiquitylation sites were predicted using the UbPred software (www.ubpred.org), where a score greater than 0.62 were considered ubiquitinated, b) the putative sumoylation sites were predicted using the SUMO plot platform (http://www.abgent.com/sumoplot), where only high probability motifs with a score of 0.5 were considered sumoylated, c) the GPS 3.0 platform was used to predict (http://gps.biocuckoo.org/) putative phosphorylation sites, implementing a highlevel threshold, and d) the webserver iPTM-mLys (http://www.jci-bioinfo.cn/iPTM-mLys) was used to predict the identifying lysine PTM sites. The online database resource Search Tool for the Retrieval of Interacting Genes (STRING) was used to identify the interactions of G6PD protein with other corresponding proteins [18].

3. RESULTS AND DISCUSSION

SNPs data

A total number of 305 SNPs (Supplementary Table 1) were retrieved from the examined data base. Out of the 305 SNPs, the 56 (18.36%) were located in the upstream gene region, 2 (0.65%) of them at the 5' prime UTR region, 170 (55.74%) at intronic regions, 6 (1.97%) were nsSNPs, 14 (4.59%) occurred synonymous, 53 (17.38%) occurred at the downstream region of the studied gene and the remaining 4 SNPs (1.31%) were other types (Figure 1).

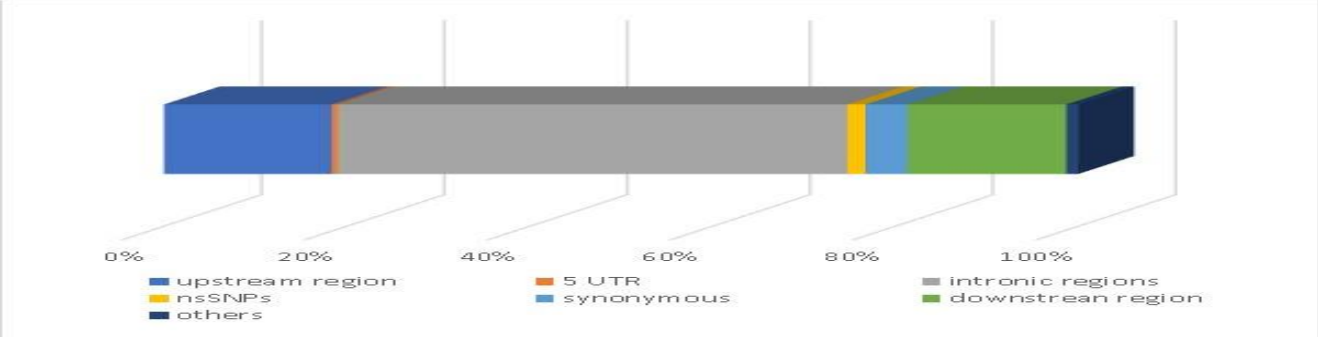

Figure 1: Schematic chart showing the categories of SNPs observed in ovine G6PD gene

Deleterious SNPs

All nsSNPs were evaluated for the observed mutation in order to detect whether the amino acid substitution may have a phenotypic effect and be damaging or not. Results are show in Table 1.

Table 1: Deleterious nsSNPs as highlighted by computational method approach

Position	Variant ID	Alleles	Residues	Codons Position	Codons	sift_class	SIFT score	Polyphen-2 Result	Polyphen-2 SCORE	snSNPs Analyzer	snSNPs Analyzer (Area Buried)
136	rs1086951364	C/T	R, C	1	CGC, TGC	deleterious	0.02	PROBAPLY DAMAGING	0.867	Disease	0.877
194	rs1094688666	C/G	D, E	3	GAC, GAG	tolerated	0.83	BENIGN	0	Neutral	0.235
196	rs1090023855	T/C	I, T	2	ATC, ACC	tolerated	0.2	BENIGN	0.535	Disease	0.671
396	rs1094240080	C/T	P, L	2	CCC, CTC	deleterious	0	PROBAPLY DAMAGING	0.998	Neutral	0.286
489	rs588075776	C/T	P, S	1	CCG, TCG	deleterious	0.03	PROBAPLY DAMAGING	0.973	Disease	0.517
495	rs429578030	A/G	K, R	2	AAA, AGA	tolerated	0.36	BENIGN	0	Neutral	0.132

From the six nsSNPS that were observed, two of them were found to be deleterious (damaging) with a tolerance index score ≤0.05. One of them (rs1094240080) showed a highly deleterious tolerance index score of 0.00. The rest were found to have a tolerant effect with an index ≥ 0.05.

Structure and stability analysis

Changes on the stability of the protein are shown in Table 2. Results revealed that all the identified

damaging nsSNPs decrease the stability of G6PD protein. The changes caused by the deleterious SNPs in the protein structure are depicted in Figure 2. According to HOPE results, for the rs1086951364 SNP a change of Arginine into a Cysteine at position 136 results in a smaller amino acid, neutral instead of positive charged and more hydrophobic than the wild-type residue. In addition, the mutation found to be located in NAD-Binding domain of G6PD. Similarly, in the rs1094240080 SNP a Proline amino acid in mutated into a Leucine at position 396. The mutant residue is bigger than the wild-type residue and is located within the Glucose-6-Phosphate Dehydrogenase C-Terminal domain. As far as it concerns the third deleterious SNP (rs588075776), a Proline is substituted by a Serine at position 489. The mutation is located in the Glucose-6-Phosphate Dehydrogenase C-Terminal domain and the mutant residue is smaller and less hydrophobic than the wild type residue.

Table 2: Ovine G6PD protein structural stability in relation with deleterious nsSNPs

Position	ID	Mutation	pH	Temperature	Stability	DDG(kcal/mol) *	RI	Characterization
136	rs1086951364	R→C	7.0	25 °C	Decrease	-0.91	4	Disease
396	rs1094240080	P→L	7.0	25 °C	Decrease	-0.42	1	Disease
489	rs588075776	P→S	7.0	25 °C	Decrease	-1.90	9	Disease

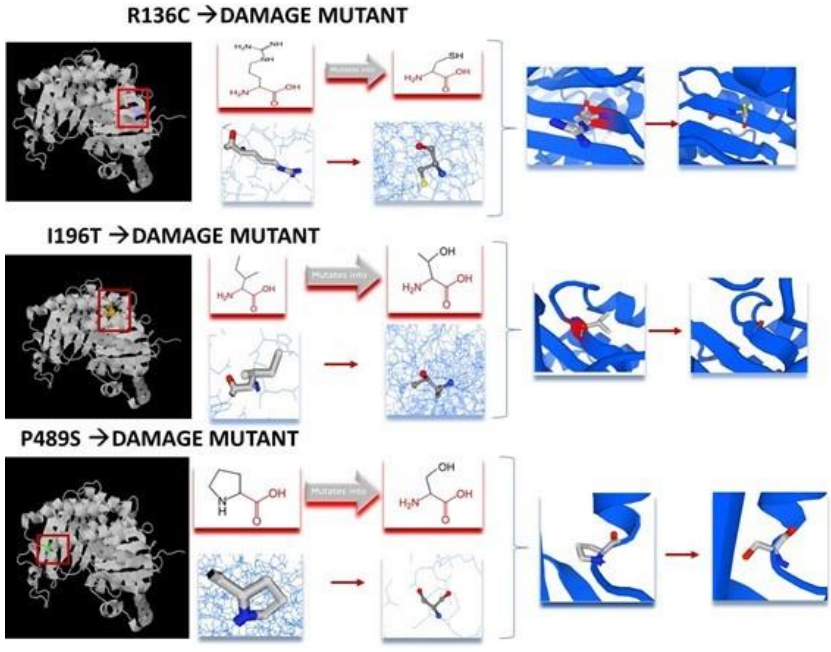

Figure 2: Structural changes caused by deleterious nsSNPs on ovine G6PD protein

Conservation Pattern of deleterious nsSNPs, post translational modification sites and protein-protein interactions

Results concerning the conservation scale of the ovine's G6PD protein amino acids are shown at Supplementary Figure 1. As far as it concerns the sites where the deleterious snSNPs were noted, two functionally and exposed residues at position 396 (respective SNP: rs1094240080) and 489 (respective SNP: rs588075776) were predicted having a highly conserved score, while the residue at position 136 (respective SNP: rs1086951364) were detected as exposed and of average conservative scale amino acid (Table 3). In addition, post translational modifications are shown in Supplementary Table 2. Many sites that might undergo phosphorylation, putative sumoylation, ubiquitylation and acetylation sites were predicted. However, sites where deleterious single nucleotide mutations were observed, did not show any potential chance for a post translational modification. As far as it concerns protein-protein interactions, Figure 3 shows functional interaction patterns of a ruminant G6PD protein to other proteins in a cell. Strong functional associations of the G6PD protein have been observed with PGLS, GPI, H6PD, PKLR, TPI1, PKM, PGD, PGM1, GCK and ENO1.

Table 3: Conservation profile of deleterious nsSNPs observed on ovine G6PD gene

nsSNP ID	Position	Conservation Score	Conservation characteristics
rs1086951364	R136	6	Middle conserved, exposed residue
rs1094240080	P396	9	Functional, highly conserved, exposed residue
rs588075776	P489	9	Functional, highly conserved, exposed residue

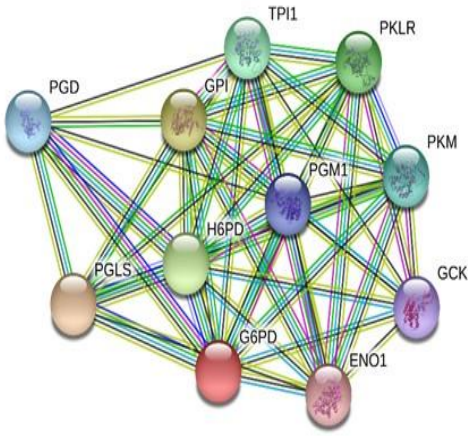

Figure 3: Protein-Protein interaction network of the ruminant G6PD molecule

DISCUSSION

The gene encoding G6PD has been widely studied in many eutherian species including human, rodents and ruminants. It is considered as a highly conserved gene, linked to the X chromosome. In human over 150 single nucleotide substitutions have been observed resulting in most of the cases in deficiency cases [10]. Although the ovine G6PD gene has been characterized and previous studies have reported a diminish of the respective enzyme activity in live animals, it remains unclear if any association of such cases with any single nucleotide change of gene's sequence exists. In the present study we report the prediction of three deleterious (damaging) SNPs (rs1086951364; rs1094240080; rs588075776), which according to the used software tools showed a high scored of occurrence. In two of them (rs1094240080, rs588075776), the wild type amino acid Proline were noted to be changed either with Leucine or Serine. The amino acid Proline is reports to create an inflexible protein pattern, forcing the whole backbone in a particular structure. Any interruption of this structure caused by a substitution of Proline with other residue interrupts this specific conformation leading to a disturb of the local structure. In addition, in the case of a bigger residue substitution compared to that of mutant it may also lead to bumps [18, 19,20]. As far as it concerns the mutation at position 136 (rs1086951364), the change of the positive charge to neutral may cause loss of interactions with other molecules or residues. In addition, as the mutant residue is more hydrophobic, and the mutation is located into the NAD-binding domain, it can result in loss of hydrogen bonds and/or disturbance of the correct folding. Results revealed also that all deleterious nsSNPs decreased G6PD protein's stability (Table 2). Hence these substitutions might cause maximum damage to the protein by affecting its stability, as according to previous reports, decreased protein stability leads to an increase of degradation, misfolding and aggregation [21]. In addition, evolutionary information is important to pinpoint amino acid substitutions that may be damaging by means of biochemical disorders or pathway deviations. The higher conserved is a functional amino acid, the more important for the function or stability of the protein is considered. Regarding the sites where the three damaging SNPs were observed, two out of the three amino acid residues were found to be not only highly conserved but also to belong to the functional ones, while the third one (rs1086951364) was noted as an exposed residue but with an average conservation rate. Arshad et al. [22] reported that the substitutions of nsSNPs which are observed in conserved regions are considered tremendously harming to a protein structure when compared to those at non-conserved locus. Thus, these both observed nsSNPs (rs1094240080; rs588075776) are rendered by highly possibility suspicious for causing further protein damage's effect. Amino acid mutations apart from the effect that may cause to the protein function, stability or protein folding, they may also affect protein–protein interactions [23]. Herein, we report the interaction of a ruminant G6PD protein (Figure 2) with other biochemical molecules. We used the bovine simulation model firstly because *Bos indicus* belongs to the same taxa with sheep and not any other corresponding ovine model has

been up to date reported, and secondly due to the very high identity score (97%) of the bovine G6PD protein (Q7YS37) that has been noted with the respective ovine (ABD34655) protein [5]. Protein interactions were observed with 6-phosphogluconolactonase, Glucose-6-phosphate isomerase, Hexose-6-Phosphate Dehydrogenase, Pyruvate kinase isozymes R, Triosephosphate isomerase, Pyruvate kinase isozymes, 6-phosphogluconate dehydrogenase, Phosphoglucomutase-1, Glucokinase and Alpha-enolase. It is obvious through that network that any change affecting protein function of G6PD can result to direct or indirect changes also to other protein's enzymatic activity or metabolic misfunctions. We further investigate the role of Post Translational Modifications (PTMs), as such modifications have been reported to play pivotal role in protein's interactions with other molecules, stability, activity and localization [24]. Phosphorylation of proteins is considered as a regulatory mechanism helping molecules to perform various functions like structural changes in protein fold, during signal transduction pathways, leading to activate or deactivate the proper proteins [25, 26]. The Small ubiquitin□related modifier (SUMO) is a protein moiety that interacts with lysine amino acid in many proteins. The addition of such particle can influence proteins' ability to interact with other molecules, alter their co-factors and control their stability. In addition, ubiquitylation plays a degradation role for proteins and assist in DNA damage repair [27, 28, 29], while acetylation, where a acetyl group is transferred to a specific site of the amino acid chain, may force regulatory modification in multiple biochemical pathways [30]. Many phosphorylation, ubiquitylation, putative sumoylation and acetylation sites were noted (Supplementary Table 2). However, such modifications were not observed in the locations of the identified deleterious nsSNPs suggesting that PTMs do not crucially influence the stability or structure of the protein in these locations. As in the respective human G6PD protein many of such nsSNPs lead to diseases cases reflecting to a diminish of the respective enzymatic activity (deficiency), we further investigate if any association of the observed deleterious ovine nsSNPs resembles with reported human cases. Interestingly, the ovine rs1086951364 nsSNPs resembles with the respective human rs7822056 nsSNP (CM970547), where at the same position (136) of the protein the same amino acid substitution is observed (R→C) in both organisms. Similarly, at the same position with the ovine rs1094240080 nsSNP (396P→L), in human counterpart a mutation is observed (CM940798; P→L), with the same residue's substitutions. Lastly at position 489, where a deleterious ovine nsSNP was observed (rs588075776; P→S) at the human counterpart a respective similar nsSNP was noted (CM077183; P→S). According to The Human Gene Mutation Database [31], all the aforementioned SNPs are related with haemoelytic anemia and G6PD deficient cases. In addition, the first SNP (position 136) has been also associated to large intensive tumors. As G6PD function is highly conserved in almost all organism, it is suggested that the noted deleterious ovine nsSNPs may be related in a great extent with non-normal functions (i.e. at enzymatic level) in sheep, as it has noted previous by other authors [11,12,13].

4. CONCLUSION

In conclusion, three out of the six observed nsSNPs in ovine G6PD were found to disturb protein fold leading to structural changes and stability protein decrease as predicted to the present study. Additionally, all of them found to have similarities with respective human disease cases, suggesting that in sheep such substitutions play a crucial role in further metabolic pathways.The results are considered of outmost importance as a first step, not only for further population genetic studies among the various ovine pure breeds or evolutionary studies but also as a first step to further elucidate their impact on biochemical lever and to the precise development of "disease" animal as a model for further veterinary or clinical (drug) studies.

CONFLICT OF INTEREST

Nothing to disclose

Supplementary Material (below Reference Section)

Supplementary Table 1: Categories, characteristics of the SNPs noted on ovine G6PD gene.

Supplementary Table 2: Post Transcriptional Modifications (PTMs) predicted on ovine G6PD protein.

Supplementary Figure 1: Predicted conservation profile of the ovine G6PD protein residues.

REFERENCES

1. Coelho M, Oliveira T, Fernandes R. Biochemistry of adipose tissue: an endocrine organ. Arch Med Sci.2013; 9(2): 191–200.
2. Vernon RG. Lipid metabolism in the adipose tissue of ruminant animals. In: Christie, W.W. (Ed.), Lipid Metabolism in Ruminant Animal. 1981, Pergamon.
3. Laliotis GP, Bizelis I, Rogdakis E. Comparative approach of the de novo fatty acid synthesis (Lipogenesis) between ruminant and non-ruminant mammalian species: From biochemical level to the main regulatory lipogenic genes. Cur Genomics. 2010; 11:168–183.
4. Laliotis GP, Argyrokastritis A, Bizelis J, Rogdakis E. Cloning, characterization of an alternative transcript of ovine glucose 6-phosphate dehydrogenase gene: comparative approach between ruminant, non-ruminant species. Gene. 2007; 388: 93-101.
5. Laliotis GP, Bizelis, J, Argyrokastritis A, Rogdakis E. Cloning, characterization, computational analysis of the 5' regulatory region of ovine glucose 6-phosphate dehydrogenase gene. Comp Bioch. Phy Part B. 2007; 147: 627-634.
6. Jiang Y, Xie M, Chen W, Talbot R, Maddox JF, Faraut T, … Dalrymple BP. The Sheep Genome Illuminates Biology of the Rumen and Lipid Metabolism. Science. 2014; 344(6188):1168–1173.
7. Loewe L. Genetic mutation. Nature Edu, 2008; 1(1):113.
8. Uppu S, Krishna A, Gopalan RP. A Review on Methods for Detecting SNP Interactions in High-Dimensional Genomic Data. In IEEE/ACM Transactions on Computational Biology. Bioinformatics, vol. 15, no. 2, 599-612.
9. Luzzatto L. G6PD deficiency: a polymorphism balanced by heterozygote advantage against

malaria. Lancet Haem, 2015; 2(10): e400 - e401.

10. Luzzatto L, Nannelli C, Notaro, R. Glucose-6-phosphate dehydrogenase deficiency. Hematol Oncol Clin North Am, 2016; 30(2): 373–93.
11. Maronpot RR. Erythrocyte glucose 6-phosphate dehydrogenase, glutathione deficiency in sheep. Can J Comp Med. 1972; 36: 55–60.
12. Calabrese E, Moore G, Ho SC. Low glucose 6-phosphate dehydrogenase activity in red blood cells, susceptibility to copper induced oxidative damage. Environ Res. 1980; 21: 366–372.
13. Moore GS, Calabrese EJ, Schulz E. The effect of in vivo ozone exposure to Dorset sheep, an animal model with low levels of erythrocyte glucose-6-phosphate dehydrogenase activity. J Environ Pathol Toxicol Oncol.1984; 5(4-5):71-78.
14. Capriotti E, Calabrese R, Casadio R. Predicting the insurgence of human genetic diseases associated to single point protein mutations with Support Vector Machines and evolutionary information. Bioinformatics. 2006; 22: 2729-2734.
15. Venselaar H, te Beek T, Kuipers RKP, Hekkelman ML, Vriend G. Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. BMC Bioinformatics, 2010; 11:548-558
16. Wu S, Zhang Y. MUSTER: Improving protein sequence profile-profile alignments by using multiple sources of structure information. Proteins: Structure, Function, Bioinformatics. 2008; 72: 547-556.
17. Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T, Ben-Tal N. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucl Acids Res, 2016; 44(W1): W344-50.
18. Betts MJ, & Russell RB. Amino acid properties and consequences of substitutions. In Bioinformatics for Geneticists, M.R. Barnes, I.C. Gray eds, Wiley, 2003.
19. Szklarczyk D, Morris JH Cook H, Kuhn M, Wyder S, Simonovic M, Santos A, Doncheva, NT, Roth A, Bork P, Jensen LJ, von Mering C. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucl Acids Res. 2017; 45: D362-368.
20. Yohannan S, Yang D, Faham S, Boulting G, Whitelegge J, Bowie JU. Proline substitutions are not easily accommodated in a membrane protein. J Mol Biol. 2004; 341(1):1-6.
21. Boone CD, Rasi V, Tu C, McKenna R. Structural and catalytic effects of proline substitution and surface loop deletion in the extended active site of human carbonic anhydrase II. FEBS J., 2015; 282(8): 1445-57.
22. Depuydt G, Shanmugam N, Rasulova M, Dhondt I, Braeckman BP. Increased Protein Stability and Decreased Protein Turnover in the Caenorhabditis elegans Ins/IGF-1 daf-2 Mutant. The Journals of Gerontology Series A: Biol Sci Med Sci, 2016; 71(12): 1553–1559.

23. Arshad M, Bhatti A, John P. Identification and in silico analysis of functional SNPs of human TAGAP protein: A comprehensive study. PLOS ONE, 2018; 13(1).
24. Reva B, Antipin Y, Ser C. Predicting the functional impact of protein mutations: application to cancer genomics. Nucl Acids Res. 2011; 39(17), e118.
25. Charlot C, Dubois-Pot H, Serchov T, Tourrette Y, Wasylyk B. A review of post-translational modifications and subcellular localization of Ets transcription factors: possible connection with cancer and involvement in the hypoxic response. Methods Mol Biol., 2010; 647: 3-30.
26. Ardito F, Giuliani M, Perrone D, Troiano G, Muzio LL. The crucial role of protein phosphorylation in cell signaling and its use as targeted therapy (Review). Int J Mol Med. 2017; 40(2): 271–280.
27. Verger A, Perdomo J, Crossley M. Modification with SUMO. A role in transcriptional regulation. EMBO reports. 2003; 4:137-142.
28. Nepveu-Traversy ME, Demogines A, Fricke T, Plourde MB, Riopel K, Veillette M, Berthoux L. A putative SUMO interacting motif in the B30.2/SPRY domain of rhesus macaque TRIM5α important for NF-κB/AP-1 signaling, HIV-1 restriction. Heliyon. 2016; 2(1): e00056.
29. Masoumi KC, Marfany G, Wu Y, Massoumi R. Putative role of SUMOylation in controlling the activity of deubiquitinating enzymes in cancer. Future Oncol. 2016; 12(4): 565-574.
30. Drazic A, Myklebust LM, Ree R, Arnesen T. The world of protein acetylation. B*iochimica et Biophysica Acta (BBA)* - Proteins and Proteomics, 2016; 1864 (10):1372-1401.
31. Stenson PD, Mort M, Ball EV, Evans K, Hayden M, Heywood S, Cooper DN. The Human Gene Mutation Database: towards a comprehensive repository of inherited mutation data for medical research, genetic diagnosis, next-generation sequencing studies. Hum Gen. 2017;136(6): 665–677.

SUPPLEMENTARY FILES

Supplementary Table 1: Categories, characteristics of the SNPs noted on ovine G6PD gene

Variant ID	Chromosome location (bp)	Allele substitution	Amino Acid change	Amino Acid location	Consequence Type	Type
rs398220443	X:77094966	T/C			upstream gene variant	SNP
rs1091440487	X:77095001	G/A			upstream gene variant	SNP
rs1091118880	X:77095041	C/G			upstream gene variant	SNP
rs413504664	X:77095084	G/A			upstream gene variant	SNP
rs595222841	X:77095153	G/A			upstream gene variant	SNP
rs425839414	X:77095190	C/T			upstream gene variant	SNP
rs1087729307	X:77095239	A/G			upstream gene variant	SNP
rs404234103	X:77095446	T/C			upstream gene variant	SNP
rs411303621	X:77095500	A/G			upstream gene variant	SNP
rs426610527	X:77095514	A/G			upstream gene variant	SNP
rs405057153	X:77095655	C/T			upstream gene variant	SNP
rs416271886	X:77095674	T/C			upstream gene variant	SNP
rs427286045	X:77095812	C/G			upstream gene variant	SNP
rs405905491	X:77095827	T/A			upstream gene variant	SNP
rs604282483	X:77096024	G/A			upstream gene variant	SNP
rs1094297157	X:77096192	T/C			upstream gene variant	SNP
rs417057314	X:77096239	C/T			upstream gene variant	SNP
rs428202590	X:77096251	T/C			upstream gene variant	SNP
rs407780136	X:77096430	A/C			upstream gene variant	SNP
rs414872558	X:77096483	T/C			upstream gene variant	SNP
rs1089491069	X:77096484-77096485	AC/-			upstream gene variant	deletion
rs430249314	X:77096530	G/A			upstream gene variant	SNP
rs1086607853	X:77096601	G/A			upstream gene variant	SNP
rs1088395730	X:77096652	G/A			upstream gene variant	SNP
rs408626798	X:77096660	C/T			upstream gene variant	SNP
rs419912153	X:77096760	T/C			upstream gene variant	SNP
rs1091099306	X:77096798	G/A			upstream gene variant	SNP
rs398309655	X:77096962	C/A			upstream gene variant	SNP
rs409511514	X:77097060	A/G			upstream gene variant	SNP
rs420495122	X:77097197	C/T			upstream gene variant	SNP

rs1088736514	X:77097202	G/T			upstream gene variant	SNP
rs399180555	X:77097747	C/A			upstream gene variant	SNP
rs1087534870	X:77097936	A/G			upstream gene variant	SNP
rs411383258	X:77097937	C/T			upstream gene variant	SNP
rs426692456	X:77097951	C/T			upstream gene variant	SNP
rs401052401	X:77098018	G/A			upstream gene variant	SNP
rs591464565	X:77098163	A/C			upstream gene variant	SNP
rs597208257	X:77098397	T/C			upstream gene variant	SNP
rs412224789	X:77098518	T/C			upstream gene variant	SNP
rs423526141	X:77098531	G/A			upstream gene variant	SNP
rs406008939	X:77098685	G/A			upstream gene variant	SNP
rs417207297	X:77098794	C/T			upstream gene variant	SNP
rs592437445	X:77098874	C/T			upstream gene variant	SNP
rs424129496	X:77098930	G/A			upstream gene variant	SNP
rs1086217464	X:77098985	A/T			upstream gene variant	SNP
rs606152301	X:77099112	A/C			upstream gene variant	SNP
rs402750679	X:77099150	C/T			upstream gene variant	SNP
rs414954198	X:77099197	T/C			upstream gene variant	SNP
rs430327347	X:77099263	A/G			upstream gene variant	SNP
rs408708526	X:77099408	C/T			upstream gene variant	SNP
rs415814785	X:77099801	T/G			upstream gene variant	SNP
rs427115296	X:77099861	C/A			upstream gene variant	SNP
rs161660019	X:77099885	C/T			upstream gene variant	SNP
rs161660018	X:77099894	G/C			upstream gene variant	SNP
rs588684766	X:77099897	G/A			upstream gene variant	SNP
rs1090979173	X:77100566	A/G			5 prime UTR variant	SNP
rs406346230	X:77100597	C/T			intron variant	SNP
rs422694436	X:77100803	A/G			intron variant	SNP
rs401134227	X:77100825	A/G			intron variant	SNP
rs600050604	X:77100871	A/T			intron variant	SNP
rs412346706	X:77100976	G/T			intron variant	SNP
rs161659994	X:77101030	A/G			intron variant	SNP
rs1088131786	X:77101153	C/T			5 prime UTR variant	SNP
rs161659990	X:77101243	T/C	A	29	synonymous variant	SNP
rs1086852987	X:77101249	G/A	T	31	synonymous variant	SNP

rs413157797	X:77101354	C/T			intron variant	SNP
rs424383042	X:77101369	T/C			intron variant	SNP
rs1088693345	X:77101461	T/C			intron variant	SNP
rs402832345	X:77101644	C/G			intron variant	SNP
rs595422050	X:77101657	G/A			intron variant	SNP
rs418166925	X:77101838	G/A			intron variant	SNP
rs425249813	X:77101872	C/T			intron variant	SNP
rs404692426	X:77101900	A/C			intron variant	SNP
rs415922123	X:77102121	G/C			intron variant	SNP
rs426946865	X:77102185	A/G			intron variant	SNP
rs405557930	X:77102227	G/A			intron variant	SNP
rs416720095	X:77102248	T/C			intron variant	SNP
rs590638670	X:77102277	T/C			intron variant	SNP
rs1085623114	X:77102314	G/A			intron variant	SNP
rs428054641	X:77102317	C/T			intron variant	SNP
rs601978982	X:77102362-77102363	CT/-			intron variant	deletion
rs1086761545	X:77102370	T/C			intron variant	SNP
rs591538676	X:77102420	G/T			intron variant	SNP
rs1088486756	X:77102621	A/G			intron variant	SNP
rs1092200671	X:77102799	C/T			intron variant	SNP
rs406426069	X:77102827	T/C			intron variant	SNP
rs597263451	X:77102918	G/A			intron variant	SNP
rs1089536456	X:77102949	G/A			intron variant	SNP
rs1091190813	X:77103113	C/-			intron variant	deletion
rs1092861173	X:77103119	C/T			intron variant	SNP
rs421789743	X:77103131	C/T			intron variant	SNP
rs1088827371	X:77103484	C/T			intron variant	SNP
rs428899446	X:77103522	G/A			intron variant	SNP
rs408331926	X:77103555	A/G			intron variant	SNP
rs1090395284	X:77103576	T/C			intron variant	SNP
rs419573114	X:77103587	T/A			intron variant	SNP
rs1087470167	X:77103607	G/A			intron variant	SNP
rs402053334	X:77103640	C/T			intron variant	SNP
rs413285599	X:77103695	C/T			intron variant	SNP

rs590290824	X:77103747	C/G			intron variant	SNP
rs420382014	X:77103831	C/T			intron variant	SNP
rs162286921	X:77103865	G/A			intron variant	SNP
rs603973671	X:77103926	G/A			intron variant	SNP
rs593495139	X:77103949	G/A			intron variant	SNP
rs414088486	X:77103969	T/G			intron variant	SNP
rs1088517777	X:77104316	C/G			intron variant	SNP
rs1094950667	X:77104470	G/T			intron variant	SNP
rs425329092	X:77104663	A/G			intron variant	SNP
rs1088550304	X:77104667	A/G			intron variant	SNP
rs1094068780	X:77104672	T/A			intron variant	SNP
rs1086118546	X:77104720	C/T			intron variant	SNP
rs1091393689	X:77104751	C/T			intron variant	SNP
rs1092754731	X:77104765	A/G			intron variant	SNP
rs1085439280	X:77104790	G/A			intron variant	SNP
rs1091678046	X: between 77104915 & 77104916	-/AT			intron variant	insertion
rs403758601	X:77104989	A/G			intron variant	SNP
rs1093250989	X:77105063	C/T			intron variant	SNP
rs411897530	X:77105113	C/T			intron variant	SNP
rs161659951	X:77105174	G/C			intron variant	SNP
rs161659950	X:77105218	A/G			intron variant	SNP
rs405680095	X:77105228	C/T			intron variant	SNP
rs162286916	X:77105344	G/C			intron variant	SNP
rs604826894	X: between 77105376 & 77105377	-/CAAGCA			intron variant	insertion
rs416848039	X:77105422	A/C			intron variant	SNP
rs428134544	X:77105450	A/G			intron variant	SNP
rs402424262	X:77105488	A/G			intron variant	SNP
rs417787839	X:77105591	A/G			intron variant	SNP
rs589073018	X:77105598-77105600	TCT/-			intron variant	deletion
rs1087697183	X:77105641	C/T			intron variant	SNP

rs162286913	X:77105649	A/T			intron variant	SNP
rs1088569535	X:77105666	C/T			intron variant	SNP
rs600244754	X:77105912	G/A			intron variant	SNP
rs589767042	X:77105947	C/T			intron variant	SNP
rs161659935	X:77106086	C/A			intron variant	SNP
rs419412235	X:77106196	G/A			intron variant	SNP
rs1088494759	X:77106418	T/A			intron variant	SNP
rs593157093	X:77106564	C/T			intron variant	SNP
rs1086174424	X:77106644	T/C			intron variant	SNP
rs601110132	X:77106661	G/A			intron variant	SNP
rs162286908	X:77106803	G/A			intron variant	SNP
rs161659923	X:77106830	C/T			intron variant	SNP
rs162286906	X:77106859	C/G			intron variant	SNP
rs430663830	X:77106959	G/A			intron variant	SNP
rs588393827	X:77107028	A/T			intron variant	SNP
rs409270739	X:77107233	C/A			intron variant	SNP
rs602067796	X:77107896	C/T			intron variant	SNP
rs1093766939	X:77107922	T/C			intron variant	SNP
rs420461836	X:77107940	A/G			intron variant	SNP
rs398919540	X:77107982	A/G			intron variant	SNP
rs414168854	X:77108033	G/A			intron variant	SNP
rs161659899	X:77108087	A/G			intron variant	SNP
rs421329185	X:77108096	C/G			intron variant	SNP
rs399772706	X:77108191	G/C			intron variant	SNP
rs589445302	X:77108276	G/A			intron variant	SNP
rs1092762593	X:77108294	A/G			intron variant	SNP
rs161659895	X:77108295	G/A			intron variant	SNP
rs1094557955	X:77108364	G/A			intron variant	SNP
rs1089814549	X:77108409	C/T			intron variant	SNP
rs423018126	X:77108444	G/C			intron variant	SNP
rs161659889	X:77108683	T/C	S	84	synonymous variant	SNP
rs412846379	X:77108737	A/G			intron variant	SNP
rs424042331	X:77108769	G/A			intron variant	SNP
rs1091858437	X:77109198	G/A	Q	111	synonymous variant	SNP
rs1086951364	X:77109271	C/T	R/C	136	missense variant	SNP

rs402504159	X:77109374	C/A			intron variant	SNP
rs417852924	X:77109403	C/G			intron variant	SNP
rs425009452	X:77109417	T/C			intron variant	SNP
rs403341290	X:77109459	A/G			intron variant	SNP
rs414558740	X:77109514	G/A			intron variant	SNP
rs429768381	X:77109567	A/T			intron variant	SNP
rs1090342023	X:77109659	C/A			intron variant	SNP
rs1087254603	X:77109718	G/A			intron variant	SNP
rs1088002321	X:77109832	C/T			intron variant	SNP
rs409316182	X:77109854	G/A			intron variant	SNP
rs603056123	X:77109866	T/-			intron variant	deletion
rs416414090	X:77110084	C/T			intron variant	SNP
rs427638377	X:77110092	A/G			intron variant	SNP
rs603132288	X:77110117	C/T			intron variant	SNP
rs410161704	X:77110119	T/C			intron variant	SNP
rs421409170	X:77110190	G/T			intron variant	SNP
rs399817320	X:77110213	G/A			intron variant	SNP
rs406961583	X:77110290	C/T			intron variant	SNP
rs596282130	X:77110297	C/G			intron variant	SNP
rs1086033813	X:77110302	G/T			intron variant	SNP
rs1091268898	X:77110309	A/G			intron variant	SNP
rs422308137	X:77110366	C/G			intron variant	SNP
rs602126733	X:77110376	C/G			intron variant	SNP
rs1086364875	X:77110463	C/G			intron variant	SNP
rs1092654043	X:77110513	C/A	T	162	splice region variant/synonymous variant	SNP
rs1094688666	X:77110609	C/G	D/E	194	missense variant	SNP
rs1090023855	X:77110614	T/C	I/T	196	missense variant	SNP
rs161659871	X:77110624	C/T	I	199	synonymous variant	SNP
rs1092688108	X:77110630	C/T	H	201	synonymous variant	SNP
rs1089957438	X:77110642	A/G	K	205	synonymous variant	SNP
rs598374520	X:77110678	C/T			splice region variant/intron variant	SNP
rs412927339	X:77110681	G/A			intron variant	SNP
rs424107654	X:77110707	A/G			intron variant	SNP

rs398494156	X:77110893	A/G			intron variant	SNP
rs413772879	X:77111013	T/C			intron variant	SNP
rs424822470	X:77111064	G/A			intron variant	SNP
rs1088913214	X:77111241	G/A	V	271	synonymous variant	SNP
rs403429672	X:77111304	T/C			intron variant	SNP
rs414632878	X:77111325	T/C			intron variant	SNP
rs425935508	X:77111408	A/G			intron variant	SNP
rs404286050	X:77111466	G/A			intron variant	SNP
rs416493711	X:77111471	G/A			intron variant	SNP
rs1090639635	X:77111485	G/A			intron variant	SNP
rs427787979	X:77111519	T/C			intron variant	SNP
rs410242949	X:77111546	T/C			intron variant	SNP
rs417402653	X:77111649	T/C	V	289	splice region variant/synonymous variant	SNP
rs1087850933	X:77111715	C/T	N	311	synonymous variant	SNP
rs161659852	X:77111784	G/A	T	334	synonymous variant	SNP
rs407044270	X:77111787	C/T	A	335	synonymous variant	SNP
rs422128189	X:77111853	T/C			intron variant	SNP
rs597045086	X:77111902	C/T			intron variant	SNP
rs400765102	X:77111914	T/C			intron variant	SNP
rs407888249	X:77112005	C/T	D	375	synonymous variant	SNP
rs1094240080	X:77112067	C/T	P/L	396	missense variant	SNP
rs161659844	X:77112077	T/C	A	399	synonymous variant	SNP
rs1085390826	X:77112215	G/A			intron variant	SNP
rs1092702608	X:77112254	G/A			splice region variant/intron variant	SNP
rs1093883329	X:77112308	C/T	C	446	synonymous variant	SNP
rs413852721	X:77113049	T/C			intron variant	SNP
rs425168463	X:77113069	C/T			intron variant	SNP
rs593208068	X:77113090	G/A			intron variant	SNP
rs1085897289	X:77113107	C/T			intron variant	SNP
rs1089606464	X:77113206	C/-			intron variant	deletion
rs1090266586	X:77113232	G/T			intron variant	SNP
rs399422902	X:77113337	C/T			intron variant	SNP

rs1087347449	X: between 77113341 & 77113342	-/A			intron variant	insertion
rs410627697	X:77113367	G/C			intron variant	SNP
rs601192292	X:77113373	A/G			intron variant	SNP
rs425746074	X:77113390	T/C			intron variant	SNP
rs588443043	X:77113446	C/T			intron variant	SNP
rs404365978	X:77113653	G/A			intron variant	SNP
rs1086829371	X:77113675	C/A			intron variant	SNP
rs411453126	X:77113780	G/A			intron variant	SNP
rs1088857455	X:77113834	C/T			intron variant	SNP
rs422706630	X:77113846	C/T			intron variant	SNP
rs1095026737	X:77113916	A/T			intron variant	SNP
rs1085853984	X:77113933	C/A			intron variant	SNP
rs1088763346	X:77113944	T/G			intron variant	SNP
rs1094099545	X:77113967	A/G			intron variant	SNP
rs406238556	X:77113989	C/T			intron variant	SNP
rs597826633	X:77114035	G/A			intron variant	SNP
rs417408099	X:77114069	A/G			intron variant	SNP
rs428524718	X:77114179	G/A			intron variant	SNP
rs403020430	X:77114183	G/C			intron variant	SNP
rs1085736537	X:77114203	C/T			intron variant	SNP
rs418378612	X:77114227	T/C			intron variant	SNP
rs588075776	X:77114243	C/T	P/S	489	missense variant	SNP
rs429578030	X:77114262	A/G	K/R	495	missense variant	SNP
rs1089500075	X:77114368	G/A			downstream gene variant	SNP
rs407970019	X:77114434	G/T			downstream gene variant	SNP
rs419218440	X:77114494	A/G			downstream gene variant	SNP
rs1090288394	X:77114523-77114524	AG/-			downstream gene variant	deletion
rs601457288	X: between 77114527 & 77114528	-/GTCCT			downstream gene variant	insertion
rs430445030	X:77114556	T/C			downstream gene variant	SNP
rs1093238195	X:77114619	G/A			downstream gene variant	SNP

rs409841097	X:77114731	A/G			downstream gene variant	SNP
rs420853257	X:77114784	A/G			downstream gene variant	SNP
rs399497166	X:77114812	A/G			downstream gene variant	SNP
rs1093693092	X:77115236-77115248	ATGAAGGACAGAC/-			downstream gene variant	deletion
rs410698591	X:77115323	C/T			downstream gene variant	SNP
rs421986311	X:77115386	A/G			downstream gene variant	SNP
rs400360645	X:77115567	T/G			downstream gene variant	SNP
rs1090657639	X:77115570	A/G			downstream gene variant	SNP
rs411537273	X:77115584	C/T			downstream gene variant	SNP
rs422780474	X:77115591	G/C			downstream gene variant	SNP
rs405299043	X:77115678	T/C			downstream gene variant	SNP
rs596060472	X:77115993	A/T			downstream gene variant	SNP
rs413455302	X:77116156	T/C			downstream gene variant	SNP
rs424462982	X:77116166	C/T			downstream gene variant	SNP
rs1088914107	X:77116196	G/A			downstream gene variant	SNP
rs403114797	X:77116303	C/T			downstream gene variant	SNP
rs418201267	X:77116368	C/A			downstream gene variant	SNP
rs429671903	X:77116394	G/A			downstream gene variant	SNP
rs403975456	X:77116629	A/G			downstream gene variant	SNP
rs415166502	X:77116661	A/G			downstream gene variant	SNP
rs591444042	X:77117058	G/A			downstream gene variant	SNP
rs1088749805	X:77117138	C/T			downstream gene variant	SNP
rs1094245349	X:77117406	G/A			downstream gene variant	SNP
rs1086472451	X:77117425	A/G			downstream gene variant	SNP
rs1088265293	X:77117431	G/A			downstream gene variant	SNP
rs430519654	X:77117459	G/A			downstream gene variant	SNP
rs408915328	X:77117478	G/A			downstream gene variant	SNP
rs421216026	X:77117481	C/A			downstream gene variant	SNP
rs1085639207	X:77117493	A/G			downstream gene variant	SNP
rs597119272	X:77117554	C/T			downstream gene variant	SNP
rs428148223	X:77117557	A/G			downstream gene variant	SNP
rs406693604	X:77117813	G/A			downstream gene variant	SNP
rs1093707033	X:77117909	C/T			downstream gene variant	SNP
rs1086245320	X:77117944	G/T			downstream gene variant	SNP

rs1091373941	X:77117960	C/T			downstream gene variant	SNP
rs1093343796	X:77118338	C/T			downstream gene variant	SNP
rs1089108413	X:77118357	C/T			downstream gene variant	SNP
rs602794889	X:77118468	G/A			downstream gene variant	SNP
rs1093542021	X:77118492	G/A			downstream gene variant	SNP
rs421997824	X:77118520	G/A			downstream gene variant	SNP
rs592335102	X:77118771	T/-			downstream gene variant	deletion
rs1090770320	X:77118892	G/A			downstream gene variant	SNP
rs598052937	X:77119014	G/-			downstream gene variant	deletion
rs594626372	X: between 77119103 & 77119104	-/GGCTCCTGGATCGTC			downstream gene variant	insertion
rs1090519829	X:77119254	G/A			downstream gene variant	SNP
rs1087601576	X:77119290	A/T			downstream gene variant	SNP

Supplementary Table 2: Post Transcriptional Modifications (PTMs) predicted on ovine G6PD protein.

Ubiquitination Sites			
Position	**Peptide**	**Score**	**Threshold**
45	GASGDLAKKKIYPTI	0.62	0.3
46	ASGDLAKKKIYPTIW	1.51	0.3
95	FKATPEEKSKLEEFF	0.93	0.3
97	ATPEEKSKLEEFFAR	0.39	0.3
175	IVEKPFGKDLQSSNQ	0.7	0.3
238	DCVILTFKEPFGTEG	0.58	0.3
275	LCLVAMEKPASTDSD	1.8	0.3
288	SDDVRDEKVKVLKCI	0.51	0.3
290	DVRDEKVKVLKCISE	0.32	0.3
320	NEEGEATKGYLDDPT	1.39	0.3
360	PFILRCGKALNERKA	1.87	0.3
366	GKALNERKAEVRLQF	0.94	0.3
403	PNEAVYTKMMTKKPG	0.76	0.3
432	GNRYKDVKLPDAYER	0.72	0.3
Putative Sumoylation Sites			

Position	Peptide	Score	Cutoff	P-value	Type
97	ATPEEKSKLEEFFAR	3.775	3.32	0.051	Sumoylation Nonconcensus
441 - 445	PDAYERLILDVFCGSQMHI	35.53	29.92	0.145	SUMO Interaction

Lysine PTM sites

Location	Prediction site
47	Acetylation
89	Acetylation
95	Acetylation
97	Acetylation
152	Acetylation
171	Acetylation
175	Acetylation
205	Acetylation
238	Acetylation
275	Acetylation
288	Acetylation
290	Acetylation
293	Acetylation
320	Acetylation
360	Acetylation
403	Acetylation
407	Acetylation
408	Acetylation
429	Acetylation
432	Acetylation
476	Acetylation
497	Acetylation
508	Acetylation
514	Acetylation

Putative Phosphorylation Sites

Position	Peptide	Kinase	Score	Cutoff
8	MAEQVALSRTQVCGI	CAMK/CAMKL/QIK/SIK1	6.312	4.074
8	MAEQVALSRTQVCGI	CAMK/MAPKAPK/MNK/MNK2	4	4

	8	MAEQVALSRTQVCGI	CAMK/PHK	5.423	3.785
	8	MAEQVALSRTQVCGI	CK1/VRK	4.3	0.977
	8	MAEQVALSRTQVCGI	CK1/VRK/VRK2	19.75	5.225
	8	MAEQVALSRTQVCGI	Other/PEK/GCN2/GCN2	7	6.95
	8	MAEQVALSRTQVCGI	Other/PEK/GCN2	6.5	5.462
	8	MAEQVALSRTQVCGI	Other/PEK/HRI/EIF2AK1	6.333	4.658
	8	MAEQVALSRTQVCGI	Other/PEK/HRI	5.5	4.792
	8	MAEQVALSRTQVCGI	Other/PEK/PEK/EIF2AK3	13	8
	8	MAEQVALSRTQVCGI	Other/PEK/PEK	13	8
	8	MAEQVALSRTQVCGI	Other/PEK/PKR	5.333	5.013
	8	MAEQVALSRTQVCGI	TKL/IRAK/IRAK4	7.2	6.78
	8	MAEQVALSRTQVCGI	TKL/MLK/MLK/MLK3	5	3.6
	8	MAEQVALSRTQVCGI	TKL/MLK/MLK	2.818	1.5
	10	EQVALSRTQVCGILR	Atypical/PIKK/ATR/ATR	9.432	8.652
	10	EQVALSRTQVCGILR	CAMK/CAMKL/MELK/MELK	0.944	0.871
	10	EQVALSRTQVCGILR	CK1/VRK	2.3	0.977
	10	EQVALSRTQVCGILR	CK1/VRK/VRK2	17.5	5.225
	10	EQVALSRTQVCGILR	Other/Haspin/Haspin	11	9
	10	EQVALSRTQVCGILR	Other/Haspin	11	9
	21	GILREELYQGDDFHQ	TK/DDR/DDR2	5.667	2.333
	21	GILREELYQGDDFHQ	TK/DDR	5.667	2.333
	21	GILREELYQGDDFHQ	TK/Src/SrcB/LCK	7.893	7.883
	31	DDFHQADTHIFIIMG	AGC/DMPK/CRIK/CIT	7.5	5.617
	31	DDFHQADTHIFIIMG	AGC/DMPK/CRIK	7.5	5.617
	31	DDFHQADTHIFIIMG	STE/STE20/FRAY/STK39	7.222	6.35
	31	DDFHQADTHIFIIMG	STE/STE20/TAO	10	9
	31	DDFHQADTHIFIIMG	STE/STE20/TAO/TAOK1	10	9
	40	IFIIMGASGDLAKKK	CAMK/MAPKAPK/MNK/MNK2	4	4
	40	IFIIMGASGDLAKKK	CAMK/RAD53/DUN1	8	7.333
	40	IFIIMGASGDLAKKK	Other/PLK/PLK2/PLK2	6.44	6.272
	40	IFIIMGASGDLAKKK	STE/STE20/MSN	4.5	3.05

	40	IFIIMGASGDLAKKK	TKL/LISK/LIMK/LIMK1	14	9
	40	IFIIMGASGDLAKKK	TKL/LISK/LIMK	14	9
	40	IFIIMGASGDLAKKK	TKL/LISK	6.25	5.388
	49	DLAKKKIYPTIWWLF	TK/Src/SrcB/BLK	9	8.462
	51	AKKKIYPTIWWLFRD	CAMK/CAMKL/PASK/PASK	10.333	7.492
	51	AKKKIYPTIWWLFRD	CAMK/CAMKL/PASK	6.5	5.767
	51	AKKKIYPTIWWLFRD	STE/STE7	0.452	0.289
	65	DGLLPEDTYIVGYAR	AGC/PDK1/PKH1	7.667	6.883
	65	DGLLPEDTYIVGYAR	AGC/PDK1/PKH2	10.333	8.317
	65	DGLLPEDTYIVGYAR	Atypical/PIKK/FRAP/TOR2	6.143	4.986
	65	DGLLPEDTYIVGYAR	CK1/CK1/CK1-D/CSNK1D	16.85	11.703
	65	DGLLPEDTYIVGYAR	Other/PLK/PLK1/PLK1	4.011	3.783
	65	DGLLPEDTYIVGYAR	STE/STE20/FRAY/STK39	8.667	6.35
	65	DGLLPEDTYIVGYAR	STE/STE20/FRAY/STLK3	11.333	5.467
	65	DGLLPEDTYIVGYAR	STE/STE20/TAO	14	9
	65	DGLLPEDTYIVGYAR	STE/STE20/TAO/TAOK1	14	9
	65	DGLLPEDTYIVGYAR	STE/STE20/YSK	9	8.95
	66	GLLPEDTYIVGYARS	TK/PDGFR/PDGFRB	6.071	5.54
	66	GLLPEDTYIVGYARS	TK/Src/SrcA/FGR	11.5	9.9
	66	GLLPEDTYIVGYARS	TK/Trk/TRKB	7.9	6.753
	66	GLLPEDTYIVGYARS	TK/Trk	8	7.239
	73	YIVGYARSRLTVADI	CAMK/MAPKAPK/MNK/MNK2	4	4
	73	YIVGYARSRLTVADI	CMGC/CDK/CDK4	6.737	6.215
	73	YIVGYARSRLTVADI	TKL/LISK	8	5.388
	76	GYARSRLTVADIRKQ	CAMK/MLCK	2	1.779
	76	GYARSRLTVADIRKQ	CAMK/PKD/PRKD2	8.5	7.292
	76	GYARSRLTVADIRKQ	CK1/CK1/CK1-D/CSNK1D	12.3	11.703
	76	GYARSRLTVADIRKQ	CK1/CK1/CK1-D	5.083	4.344
	76	GYARSRLTVADIRKQ	Other/AUR/IPL1	9	6.994
	76	GYARSRLTVADIRKQ	TKL/IRAK/IRAK4	7.4	6.78
	84	VADIRKQSEPFFKAT	AGC/NDR/LATS/LATS1	8.8	6.9
	84	VADIRKQSEPFFKAT	AGC/NDR/LATS/LATS2	9.571	7.171
	84	VADIRKQSEPFFKAT	AGC/NDR/LATS	9.5	7.175

	84	VADIRKQSEPFFKAT	AGC/NDR	7.1	5.725
	84	VADIRKQSEPFFKAT	AGC/PKA/PKACA	5.589	4.405
	84	VADIRKQSEPFFKAT	AGC/PKA/PKACB	10	6.283
	84	VADIRKQSEPFFKAT	AGC/PKA	5.381	4.604
	84	VADIRKQSEPFFKAT	AGC/PKC/PKCd	4.378	3.771
	84	VADIRKQSEPFFKAT	AGC/PKG/PRKG1	12.767	9.542
	84	VADIRKQSEPFFKAT	AGC/PKG/PRKG2	6.474	4.332
	84	VADIRKQSEPFFKAT	AGC/PKG	4.119	3.869
	84	VADIRKQSEPFFKAT	AGC/RSK/MSK/RPS6KA5	14.519	12.002
	84	VADIRKQSEPFFKAT	AGC/RSK/RSKp90/RPS6KA3	10.941	10.513
	84	VADIRKQSEPFFKAT	AGC/RSK/RSKp90	6.787	5.671
	84	VADIRKQSEPFFKAT	AGC	1.361	1.279
	84	VADIRKQSEPFFKAT	Atypical/Alpha/ChaK	5.625	3.588
	84	VADIRKQSEPFFKAT	Atypical/Alpha/ChaK/TRPM7	8.143	6.629
	84	VADIRKQSEPFFKAT	Atypical/Alpha	3.375	1.988
	84	VADIRKQSEPFFKAT	CAMK/CAMK2/CAMK2D	5.083	4.204
	84	VADIRKQSEPFFKAT	CAMK/CAMK2	3.717	3.65
	84	VADIRKQSEPFFKAT	CAMK/CAMKL/AMPK/AMPKA1	11	7.125
	84	VADIRKQSEPFFKAT	CAMK/CAMKL/CHK1/CHEK1	3.69	2.783
	84	VADIRKQSEPFFKAT	CAMK/CAMKL/MARK/MARK3	8	7.5
	84	VADIRKQSEPFFKAT	CAMK/CAMKL	9.18	8.942
	84	VADIRKQSEPFFKAT	CAMK/MAPKAPK/MK2/MAPKAPK2	5.869	5.4
	84	VADIRKQSEPFFKAT	CAMK/MAPKAPK/MK2	6.311	5.725
	84	VADIRKQSEPFFKAT	CAMK/MAPKAPK/MNK/MNK2	4	4
	84	VADIRKQSEPFFKAT	CAMK/RAD53	5.604	5.343
	84	VADIRKQSEPFFKAT	STE/STE-Unique/NIK	7.667	6.283
	84	VADIRKQSEPFFKAT	STE/STE-Unique/NIK	7.667	6.283
	84	VADIRKQSEPFFKAT	STE/STE20/PAKA/PAK1	10.289	9.701
	84	VADIRKQSEPFFKAT	STE/STE20/PAKA/PAK2	2.688	2.05
	84	VADIRKQSEPFFKAT	STE/STE20/PAKB/PAK5	7.5	6.238
	84	VADIRKQSEPFFKAT	STE/STE20/PAKB	5.2	4.79
	84	VADIRKQSEPFFKAT	TKL/RAF/RAF/RAF1	4.778	2.47
	84	VADIRKQSEPFFKAT	TKL/RAF	2.867	2.289

	91	SEPFFKATPEEKSKL	CMGC/CDK/CDC2/CDK1	14.407	8.98
	91	SEPFFKATPEEKSKL	CMGC/CDK/CDC2	10.42	8.778
	91	SEPFFKATPEEKSKL	CMGC/CDK/CDK2/CDC28	16.5	13.343
	91	SEPFFKATPEEKSKL	CMGC/CDK/CDK4/CDK6	9.667	6.758
	91	SEPFFKATPEEKSKL	CMGC/CDK/CDK4	7.737	6.215
	91	SEPFFKATPEEKSKL	CMGC/CDK/CDK5/CDK5	26.753	15.952
	91	SEPFFKATPEEKSKL	CMGC/CDK/CDK9/CDK9	9	7.718
	91	SEPFFKATPEEKSKL	CMGC/CDK	6.931	5.994
	91	SEPFFKATPEEKSKL	CMGC/CK2/CK2a1	2.903	2.848
	91	SEPFFKATPEEKSKL	CMGC/DYRK/HIPK/HIPK2	2.103	1.841
	91	SEPFFKATPEEKSKL	CMGC/GSK/GSK3A	7.176	4.488
	91	SEPFFKATPEEKSKL	CMGC/MAPK/ERK/Erk1	25.653	17.725
	91	SEPFFKATPEEKSKL	CMGC/MAPK/ERK/Erk4	11	9
	91	SEPFFKATPEEKSKL	CMGC/MAPK/p38/MAPK11	10.222	8.192
	91	SEPFFKATPEEKSKL	CMGC/MAPK/p38/MAPK14	3.494	3.342
	91	SEPFFKATPEEKSKL	CMGC	3.594	2.953
	91	SEPFFKATPEEKSKL	Other/NEK/NEK2/NEK2	5.4	5.105
	91	SEPFFKATPEEKSKL	Other/NEK/NEK2	5.4	5.105
	91	SEPFFKATPEEKSKL	STE/STE20/MST/MST1	52	9.234
	91	SEPFFKATPEEKSKL	STE/STE20	11.39	2.369
	91	SEPFFKATPEEKSKL	STE/STE7/MEK7/MAP2K7	11.167	8.815
	91	SEPFFKATPEEKSKL	STE/STE7/MEK7	12.833	10.315
	96	KATPEEKSKLEEFFA	AGC/GRK	8.414	7.991
	96	KATPEEKSKLEEFFA	AGC/PKC/PKCi/PRKCI	6.375	6.256
	96	KATPEEKSKLEEFFA	CAMK/CAMK1/MLKA	6	5.667
	96	KATPEEKSKLEEFFA	CAMK/CAMKL/MELK/MELK	0.889	0.871
	96	KATPEEKSKLEEFFA	CAMK/CAMKL/MELK	1.5	1.36
	96	KATPEEKSKLEEFFA	CAMK/MAPKAPK/MNK/MNK2	4	4
	96	KATPEEKSKLEEFFA	Other/NEK/NEK6/NEK6	8.333	5.103
	96	KATPEEKSKLEEFFA	Other/PLK/PLK2/PLK3	4.333	4.021
	96	KATPEEKSKLEEFFA	Other/PLK/PLK2	7.279	6.617
	96	KATPEEKSKLEEFFA	Other/ULK/ULK/ULK1	5.667	4.508
	107	EFFARNFYVAGQYDD	TK/EGFR/EGFR	5.882	5.442

	107	EFFARNFYVAGQYDD	TK/VEGFR/FLT4	7.273	5.414
	112	NFYVAGQYDDTASYE	TK/Axl/AXL	3.667	3.333
	112	NFYVAGQYDDTASYE	TK/Csk/CTK	12	11.4
	112	NFYVAGQYDDTASYE	TK/Eph/EphB5	2.333	2.142
	112	NFYVAGQYDDTASYE	TK/FGFR/FGFR4	4.333	2.917
	112	NFYVAGQYDDTASYE	TK/Fer	8.579	7.78
	115	VAGQYDDTASYERLN	CAMK/CAMKL/LKB	4.967	2.821
	115	VAGQYDDTASYERLN	CAMK/CAMKL/LKB/STK11	4.931	2.793
	115	VAGQYDDTASYERLN	Other/PLK/PLK1/CDC5	11.111	6.964
	115	VAGQYDDTASYERLN	Other/TTK/MPS1	9.111	7.039
	115	VAGQYDDTASYERLN	Other/TTK	6.188	5.303
	115	VAGQYDDTASYERLN	STE/STE20/PAKB/PAK5	6.25	6.238
	115	VAGQYDDTASYERLN	STE/STE20/TAO	9	9
	115	VAGQYDDTASYERLN	STE/STE20/TAO/TAOK1	9	9
	115	VAGQYDDTASYERLN	STE/STE7/MEK3/MAP2K4	2	1.325
	115	VAGQYDDTASYERLN	TKL/MLK/MLK/MLK3	5.75	3.6
	115	VAGQYDDTASYERLN	TKL/MLK/MLK	2.545	1.5
	117	GQYDDTASYERLNGH	AGC/RSK/RSKp90/RPS6KA2	7.333	6.617
	117	GQYDDTASYERLNGH	AGC/SGK/SGK3	6.143	5.871
	117	GQYDDTASYERLNGH	Atypical/PDHK/PDHK/PDK3	6.5	6.138
	117	GQYDDTASYERLNGH	CAMK/CAMKL/NuaK/NUAK1	5.375	4.519
	117	GQYDDTASYERLNGH	CAMK/CAMKL/NuaK	5.375	4.519
	117	GQYDDTASYERLNGH	CAMK/MAPKAPK/MNK/MNK2	4	4
	117	GQYDDTASYERLNGH	CMGC/CLK/CLK1	3.25	2.638
	117	GQYDDTASYERLNGH	CMGC/CLK	7.25	5.806
	118	QYDDTASYERLNGHI	TK/Eph/EphA2	9.2	4.7
	118	QYDDTASYERLNGHI	TK/Eph/EphB5	2.333	2.142
	118	QYDDTASYERLNGHI	TK/PDGFR/FLT3	1.778	1.067
	118	QYDDTASYERLNGHI	TK/Ret/RET	5.037	4.574

	118	QYDDTASYERLNGHI	TK/Src/SrcA/FGR	12.2	9.9
	118	QYDDTASYERLNGHI	TK/Src/SrcA/FYN	8.964	6.401
	118	QYDDTASYERLNGHI	TK/Src/SrcB/BLK	8.5	8.462
	118	QYDDTASYERLNGHI	TK/Src/SrcB/LYN	6.063	5.358
	118	QYDDTASYERLNGHI	TK/Trk/TRKA	24.4	18.155
	118	QYDDTASYERLNGHI	TK/VEGFR	6.519	5.372
	118	QYDDTASYERLNGHI	TK	1.417	1.036
	132	INALHRGTQTNRLFY	CMGC/CDK/CDK7	6.118	4.112
	132	INALHRGTQTNRLFY	Other/Wnk	3.917	3.514
	132	INALHRGTQTNRLFY	Other/Wnk/Wnk1	3.917	3.514
	132	INALHRGTQTNRLFY	Other/Wnk/Wnk1	3.917	3.514
	132	INALHRGTQTNRLFY	STE/STE20/PAKA	9.547	8.979
	132	INALHRGTQTNRLFY	STE/STE20	3.006	2.369
	134	ALHRGTQTNRLFYLV	CAMK/CAMK1/MLKA	6.333	5.667
	134	ALHRGTQTNRLFYLV	CAMK/CAMKL/AMPK/AMPKA1	7.25	7.125
	134	ALHRGTQTNRLFYLV	Other/PEK/HRI/EIF2AK1	4.833	4.658
	134	ALHRGTQTNRLFYLV	Other/PEK/HRI	4.833	4.792
	134	ALHRGTQTNRLFYLV	TKL/MLK/TAK1/TAK1	7	4.942
	139	TQTNRLFYLVLPPTV	TK/Axl/AXL	5.333	3.333
	145	FYLVLPPTVYEAVTK	STE/STE20/PAKA/STE20	6.8	5.67
	147	LVLPPTVYEAVTKNI	TK/Eph/EphB2	5.071	4.611
	147	LVLPPTVYEAVTKNI	TK/Fer	8.947	7.78
	147	LVLPPTVYEAVTKNI	TK/Src/SrcA/FYN	9.286	6.401
	147	LVLPPTVYEAVTKNI	TK/Src/SrcB/LCK	8.28	7.883
	147	LVLPPTVYEAVTKNI	TK/Tec/ITK	5.4	2.44
	147	LVLPPTVYEAVTKNI	TK	1.102	1.036
	157	VTKNIHETCMSQTGW	CK1/CK1/CK1-D/CK1e	7.059	6.526
	157	VTKNIHETCMSQTGW	CMGC/CDK/CDK4/CDK6	9.333	6.758
	157	VTKNIHETCMSQTGW	CMGC/CDK/CDK9	4.7	4.647
	157	VTKNIHETCMSQTGW	STE/STE7/MEK3/MAP2K3	5.778	4.309
	157	VTKNIHETCMSQTGW	STE/STE7/MEK3	2.308	2.053
	160	NIHETCMSQTGWNRI	AGC/GRK/GRK/GPRK7	5.667	5.483
	160	NIHETCMSQTGWNRI	Atypical/PIKK/ATM/ATM	3.872	3.362

160	NIHETCMSQTGWNRI	Atypical/PIKK/ATM	4.188	4.149
160	NIHETCMSQTGWNRI	Atypical/PIKK/ATR/ATR	9.455	8.652
160	NIHETCMSQTGWNRI	Atypical/PIKK/ATR/MEC1	8.667	7.333
160	NIHETCMSQTGWNRI	Atypical/PIKK/ATR	7.064	5.485
160	NIHETCMSQTGWNRI	Atypical/PIKK/DNAPK/PRKDC	6.886	5.924
160	NIHETCMSQTGWNRI	Atypical/PIKK/DNAPK	6.4	5.987
160	NIHETCMSQTGWNRI	Atypical/PIKK	6.683	6.439
160	NIHETCMSQTGWNRI	CAMK/MAPKAPK/MNK/MNK2	4	4
160	NIHETCMSQTGWNRI	Other/Bud32	8.75	7.738
160	NIHETCMSQTGWNRI	STE/NDR/NDR/CBK1	7.5	6.138
160	NIHETCMSQTGWNRI	STE/NDR/NDR	7.5	6.138
160	NIHETCMSQTGWNRI	STE/NDR	7.5	6.138
160	NIHETCMSQTGWNRI	STE/NDR/NDR/CBK1	7.5	6.138
160	NIHETCMSQTGWNRI	STE/NDR/NDR	7.5	6.138
160	NIHETCMSQTGWNRI	STE/NDR	7.5	6.138
160	NIHETCMSQTGWNRI	STE/STE20/KHS	10.25	8.4
162	HETCMSQTGWNRIIV	CAMK/CAMKL/PASK/PASK	7.5	7.492
162	HETCMSQTGWNRIIV	CAMK/CAMKL/PASK	6	5.767
162	HETCMSQTGWNRIIV	Other/NAK/BIKE/AAK1	6.333	4.9
162	HETCMSQTGWNRIIV	Other/NAK/BIKE	9	7.67
162	HETCMSQTGWNRIIV	Other/NAK	9	7.67
162	HETCMSQTGWNRIIV	Other/ULK/ULK	7.909	5.964
162	HETCMSQTGWNRIIV	STE/STE7/MEK3/MAP2K6	9.571	8.35
179	PFGKDLQSSNQLSNH	CAMK/DAPK/DAPK/DAPK1	6.3	4.846
179	PFGKDLQSSNQLSNH	CAMK/MAPKAPK/MNK/MNK2	4	4
179	PFGKDLQSSNQLSNH	CAMK/PIM	7.571	6.783
180	FGKDLQSSNQLSNHI	AGC/PKC/PKCi	11.646	10.233
180	FGKDLQSSNQLSNHI	CAMK/MAPKAPK/MNK/MNK2	4	4
180	FGKDLQSSNQLSNHI	Other/ULK/ULK/ULK1	5.333	4.508
184	LQSSNQLSNHIDYLL	CAMK/CAMK2/CAMK2B	9.667	8.277

	184	LQSSNQLSNHIDYLL	CAMK/MAPKAPK/MNK/MNK2	4	4
	184	LQSSNQLSNHIDYLL	Other/Bud32	10.75	7.738
	184	LQSSNQLSNHIDYLL	Other/PEK/HRI/EIF2AK1	5	4.658
	184	LQSSNQLSNHIDYLL	Other/PEK/PEK/EIF2AK3	8	8
	184	LQSSNQLSNHIDYLL	Other/PEK/PEK	8	8
	184	LQSSNQLSNHIDYLL	STE/STE7/MEK3/MAP2K4	2	1.325
	184	LQSSNQLSNHIDYLL	TKL/MLK/MLK	1.636	1.5
	189	QLSNHIDYLLHEDQI	TK/FGFR/FGFR3	9	8.638
	189	QLSNHIDYLLHEDQI	TK/Jak/JAK3	8	5.488
	189	QLSNHIDYLLHEDQI	TK/PDGFR/PDGFRA	12	11.527
	197	LLHEDQIYRIDHYLG	TK/Axl/MER	17.5	13.6
	197	LLHEDQIYRIDHYLG	TK/Ret/RET	4.852	4.574
	197	LLHEDQIYRIDHYLG	TK/Src/SrcA/SRC	3.051	2.269
	223	FANRIFGTIWNRDNI	AGC/DMPK/GEK/MRCKa	4.8	4.22
	223	FANRIFGTIWNRDNI	STE/STE20/MSN/MAP4K4	9	9
	223	FANRIFGTIWNRDNI	STE/STE20/MSN/MINK1	9	9
	223	FANRIFGTIWNRDNI	STE/STE20/MSN/TNIK	9	9
	236	NIDCVILTFKEPFGT	Other/WEE/Myt1/PKMYT1	8.667	5.583
	236	NIDCVILTFKEPFGT	Other/WEE/Myt1	8.667	5.583
	236	NIDCVILTFKEPFGT	Other/WEE	8.667	5.583
	243	TFKEPFGTEGRGGYF	Other/PLK/PLK2/PLK2	6.76	6.272
	243	TFKEPFGTEGRGGYF	Other/TTK	6.5	5.303
	243	TFKEPFGTEGRGGYF	Other/TTK/TTK	58.182	55.045
	249	GTEGRGGYFDEFWII	TK/Eph/EphA4	9.5	8
	278	VAMEKPASTDSDDVR	AGC/PKC/PKCd/PRKCQ	4.364	3.963
	278	VAMEKPASTDSDDVR	AGC/PKC/PKCh/PRKCE	3.516	2.939
	278	VAMEKPASTDSDDVR	AGC/PKC/PKCh	8.359	6.801
	278	VAMEKPASTDSDDVR	CAMK/CAMK-Unique	4.667	4.189
	278	VAMEKPASTDSDDVR	CAMK/CAMK2/CAMK2A	5.18	4.003
	278	VAMEKPASTDSDDVR	CAMK/CAMKL/MARK/MARK1	8.333	7.433
	278	VAMEKPASTDSDDVR	CAMK/CAMKL/NuaK/NUAK1	5.75	4.519
	278	VAMEKPASTDSDDVR	CAMK/CAMKL/NuaK	5.75	4.519

	278	VAMEKPASTDSDDVR	CAMK/MAPKAPK/MK5/MAPKAPK5	17	9.696
	278	VAMEKPASTDSDDVR	CAMK/MAPKAPK/MK5	17	9.696
	278	VAMEKPASTDSDDVR	CAMK/MAPKAPK/MNK/MNK1	11.5	10.35
	278	VAMEKPASTDSDDVR	CAMK/MAPKAPK/MNK/MNK2	4	4
	278	VAMEKPASTDSDDVR	CAMK/RAD53/DUN1	8.333	7.333
	278	VAMEKPASTDSDDVR	CK1/CK1	6.008	3.998
	278	VAMEKPASTDSDDVR	CMGC/CK2/CK2a1	4.097	2.848
	278	VAMEKPASTDSDDVR	CMGC/CLK/CLK2	4.714	4.679
	278	VAMEKPASTDSDDVR	CMGC/DYRK/DYRK1/DYRK1A	6.636	6.152
	278	VAMEKPASTDSDDVR	Other/NEK/NEK11	17	12.967
	278	VAMEKPASTDSDDVR	STE/STE20/PAKA/CLA4	9.25	6.75
	278	VAMEKPASTDSDDVR	STE/STE20/PAKA/STE20	7.6	5.67
	278	VAMEKPASTDSDDVR	TKL/STKR/STKR1	12.75	8.6
	279	AMEKPASTDSDDVRD	CAMK/CAMKL/QIK/SIK2	8.167	7.514
	279	AMEKPASTDSDDVRD	CAMK/PIM/PIM1	10	8.973
	279	AMEKPASTDSDDVRD	CK1/CK1/CK1-D/CK1e	9.941	6.526
	279	AMEKPASTDSDDVRD	CMGC/CK2/CK2a1	5.528	2.848
	279	AMEKPASTDSDDVRD	CMGC/CK2	10.138	9.894
	279	AMEKPASTDSDDVRD	Other/NEK/NEK6	2.833	2.494
	279	AMEKPASTDSDDVRD	STE/STE7	0.29	0.289
	281	EKPASTDSDDVRDEK	Atypical/TAF1	5.125	4.681
	281	EKPASTDSDDVRDEK	Atypical/TAF1/TAF1	5.125	4.681
	281	EKPASTDSDDVRDEK	CAMK/MAPKAPK/MNK/MNK2	4	4
	281	EKPASTDSDDVRDEK	CK1/CK1/CK1-D	4.861	4.344
	281	EKPASTDSDDVRDEK	CK1/TTBK	6.3	6.136
	281	EKPASTDSDDVRDEK	CMGC/CK2/CK2a1	4.062	2.848
	281	EKPASTDSDDVRDEK	CMGC/CK2	11.55	9.894
	281	EKPASTDSDDVRDEK	Other/NEK/NEK11/NEK11	7.667	5.067

281	EKPASTDSDDVRDEK	Other/NEK/NEK11	20.333	12.967
281	EKPASTDSDDVRDEK	Other/PLK/PLK2/PLK2	6.56	6.272
281	EKPASTDSDDVRDEK	Other/ULK	3.077	2.631
296	VKVLKCISEVQASNV	Atypical/TAF1	5.125	4.681
296	VKVLKCISEVQASNV	Atypical/TAF1/TAF1	5.125	4.681
296	VKVLKCISEVQASNV	CAMK/MAPKAPK/MNK/MNK2	4	4
296	VKVLKCISEVQASNV	Other/AUR/AurA	3.65	3.498
296	VKVLKCISEVQASNV	TKL/LISK	8	5.388
301	CISEVQASNVVLGQY	AGC/GRK/GRK/GPRK7	5.667	5.483
301	CISEVQASNVVLGQY	CAMK/MAPKAPK/MNK/MNK2	4	4
301	CISEVQASNVVLGQY	CAMK/PIM/PIM3	2.3	2.11
301	CISEVQASNVVLGQY	Other/Bud32	8.25	7.738
301	CISEVQASNVVLGQY	TKL/LISK	7.5	5.388
308	SNVVLGQYVGNPNEE	TK/Syk/ZAP70	5.452	4.122
308	SNVVLGQYVGNPNEE	TK/Tec	4.049	3.674
319	PNEEGEATKGYLDDP	AGC/PKC/PKCh/PRKCH	8.9	3.79
319	PNEEGEATKGYLDDP	Other/Haspin/Haspin	9	9
319	PNEEGEATKGYLDDP	Other/Haspin	9	9
319	PNEEGEATKGYLDDP	TKL/IRAK/IRAK1	6.25	2.8
319	PNEEGEATKGYLDDP	TKL/IRAK	3.75	2.088
322	EGEATKGYLDDPTVP	TK/EGFR/EGFR	6.368	5.442
322	EGEATKGYLDDPTVP	TK/Jak	4.697	4.054
327	KGYLDDPTVPRGSTT	Atypical/Alpha/ChaK	4.875	3.588
327	KGYLDDPTVPRGSTT	Atypical/PIKK/FRAP/MTOR	5.379	4.258
327	KGYLDDPTVPRGSTT	CK1/CK1/CK1-D	4.833	4.344
327	KGYLDDPTVPRGSTT	Other/PLK/PLK2	7.163	6.617
327	KGYLDDPTVPRGSTT	Other/TTK/MPS1	8.111	7.039
327	KGYLDDPTVPRGSTT	TKL/MLK	12.68	9.761
327	KGYLDDPTVPRGSTT	TKL	5	4.148
332	DPTVPRGSTTATFAA	AGC/DMPK/GEK/DMPK	10.5	6.088
332	DPTVPRGSTTATFAA	AGC/PKA	4.967	4.604
332	DPTVPRGSTTATFAA	AGC/PKG/PRKG2	4.895	4.332
332	DPTVPRGSTTATFAA	AGC/RSK/MSK/RPS6KA4	8.75	7.684

	332	DPTVPRGSTTATFAA	Atypical/Alpha/ChaK	4.5	3.588
	332	DPTVPRGSTTATFAA	Atypical/Alpha/ChaK/TRPM7	9	6.629
	332	DPTVPRGSTTATFAA	CAMK/CAMKL/BRSK/BRSK2	8.667	5.333
	332	DPTVPRGSTTATFAA	CAMK/CAMKL/MARK/MARK1	9.111	7.433
	332	DPTVPRGSTTATFAA	CAMK/CAMKL/MELK	1.667	1.36
	332	DPTVPRGSTTATFAA	CAMK/MAPKAPK/MNK/MNK1	13	10.35
	332	DPTVPRGSTTATFAA	CAMK/MAPKAPK/MNK/MNK2	4	4
	332	DPTVPRGSTTATFAA	CK1/TTBK	8.9	6.136
	332	DPTVPRGSTTATFAA	Other/NEK/NEK11	13	12.967
	332	DPTVPRGSTTATFAA	Other/Wnk	3.75	3.514
	332	DPTVPRGSTTATFAA	Other/Wnk/Wnk1	3.75	3.514
	332	DPTVPRGSTTATFAA	Other/Wnk	3.75	3.514
	332	DPTVPRGSTTATFAA	Other/Wnk/Wnk1	3.75	3.514
	333	PTVPRGSTTATFAAA	AGC/AKT/AKT1	3.721	2.98
	333	PTVPRGSTTATFAAA	AGC/DMPK/CRIK/CIT	9	5.617
	333	PTVPRGSTTATFAAA	AGC/DMPK/CRIK	9	5.617
	333	PTVPRGSTTATFAAA	AGC/NDR/LATS/LATS1	8.6	6.9
	333	PTVPRGSTTATFAAA	AGC/NDR/LATS/LATS2	8.286	7.171
	333	PTVPRGSTTATFAAA	AGC/NDR/LATS	8	7.175
	333	PTVPRGSTTATFAAA	AGC/PKG/PRKG1	10.767	9.542
	333	PTVPRGSTTATFAAA	CAMK/CAMK2/CAMK2D	4.417	4.204
	333	PTVPRGSTTATFAAA	CAMK/MAPKAPK/MK2/MAPKAPK3	1.583	1.45
	333	PTVPRGSTTATFAAA	CAMK/PHK	3.808	3.785
	333	PTVPRGSTTATFAAA	CMGC/DYRK/DYRK2	12.444	11.395
	333	PTVPRGSTTATFAAA	TKL/MLK/MLK/ZAK	10.25	9.25
	334	TVPRGSTTATFAAAV	AGC/GRK/GRK/GPRK7	6.333	5.483
	334	TVPRGSTTATFAAAV	AGC/GRK/GRK/GRK	9.556	9.35
	336	PRGSTTATFAAAVLY	AGC/PDK1/PKH1	7.667	6.883

	336	PRGSTTATFAAAVLY	AGC/PDK1/PKH2	10	8.317
	336	PRGSTTATFAAAVLY	AGC/PDK1	3.537	2.39
	336	PRGSTTATFAAAVLY	AGC/SGK/SGK3	6	5.871
	336	PRGSTTATFAAAVLY	Atypical/PDHK	8.167	4.981
	336	PRGSTTATFAAAVLY	Atypical/PIKK/FRAP/TOR2	5.429	4.986
	336	PRGSTTATFAAAVLY	CAMK/CAMKL/MELK	1.389	1.36
	336	PRGSTTATFAAAVLY	CK1/CK1/CK1-A	5.526	3.887
	336	PRGSTTATFAAAVLY	STE/STE11/ASK/MAP3K5	6.333	6.006
	336	PRGSTTATFAAAVLY	STE/STE11/ASK	6.167	6
	336	PRGSTTATFAAAVLY	STE/STE11/ASK/MAP3K5	6.333	6.006
	336	PRGSTTATFAAAVLY	STE/STE11/ASK	6.167	6
	336	PRGSTTATFAAAVLY	STE/STE20/TAO	9	9
	336	PRGSTTATFAAAVLY	STE/STE20/TAO/TAOK1	9	9
	336	PRGSTTATFAAAVLY	STE/STE20/YSK	9	8.95
	336	PRGSTTATFAAAVLY	TKL/RIPK	3.667	1.75
	343	TFAAAVLYVENERWD	TK/Ack	3.143	2.055
	343	TFAAAVLYVENERWD	TK/Csk/CSK	13.037	12.276
	343	TFAAAVLYVENERWD	TK/PDGFR/PDGFRA	15.6	11.527
	343	TFAAAVLYVENERWD	TK/Src/SRM	3.682	1.527
	343	TFAAAVLYVENERWD	TK/Tie	8.714	8.494
	401	VQPNEAVYTKMMTKK	TK/Abl	4.211	3.718
	401	VQPNEAVYTKMMTKK	TK/Axl/AXL	5	3.333
	401	VQPNEAVYTKMMTKK	TK/FGFR/FGFR1	8.04	4.789
	401	VQPNEAVYTKMMTKK	TK/Fak/FAK	8.312	7.683
	401	VQPNEAVYTKMMTKK	TK/Fer/FER	10.4	7.185
	401	VQPNEAVYTKMMTKK	TK/Fer/FEs	7.364	7.262
	401	VQPNEAVYTKMMTKK	TK/Fer	7.947	7.78
	401	VQPNEAVYTKMMTKK	TK/Jak/JAK1	8.333	5.035

	401	VQPNEAVYTKMMTKK	TK/Jak/JAK2	17.045	12.549
	401	VQPNEAVYTKMMTKK	TK/Ret/RET	5.481	4.574
	401	VQPNEAVYTKMMTKK	TK/Ret	6.593	4.572
	401	VQPNEAVYTKMMTKK	TK/Src/SRM	2.409	1.527
	401	VQPNEAVYTKMMTKK	TK/Src/SrcA/FYN	7.012	6.401
	401	VQPNEAVYTKMMTKK	TK/Src/SrcA	4.106	2.268
	401	VQPNEAVYTKMMTKK	TK/Src/SrcB/LCK	8.293	7.883
	401	VQPNEAVYTKMMTKK	TK/Src	1.72	1.63
	401	VQPNEAVYTKMMTKK	TK/Tec/TXK	7.286	6.294
	401	VQPNEAVYTKMMTKK	TK/Tie/TIE2	9.5	8.575
	401	VQPNEAVYTKMMTKK	TK/Tie	9.143	8.494
	401	VQPNEAVYTKMMTKK	TK	1.216	1.036
	402	QPNEAVYTKMMTKKP	STE/STE20/MST	25.184	9.532
	406	AVYTKMMTKKPGMFF	AGC/DMPK	2.986	2.427
	406	AVYTKMMTKKPGMFF	AGC/PKC/PKCh	7.436	6.801
	406	AVYTKMMTKKPGMFF	CAMK/CAMKL/MELK/MELK	1.278	0.871
	406	AVYTKMMTKKPGMFF	CAMK/CAMKL/MELK	1.889	1.36
	406	AVYTKMMTKKPGMFF	CAMK/RAD53/RAD53	6.667	5.467
	406	AVYTKMMTKKPGMFF	CK1/VRK/VRK1	3.9	3.159
	406	AVYTKMMTKKPGMFF	CMGC/CDK/CDK7	4.206	4.112
	406	AVYTKMMTKKPGMFF	Other/Haspin/Haspin	9	9
	406	AVYTKMMTKKPGMFF	Other/Haspin	9	9
	418	MFFNPEESELDLTYG	AGC/GRK/BARK	5.256	4.78
	418	MFFNPEESELDLTYG	CAMK/MAPKAPK/MNK/MNK2	4	4
	418	MFFNPEESELDLTYG	CMGC/CDK/CDK7/CDK7	0.727	0.406
	418	MFFNPEESELDLTYG	CMGC/CK2	13.552	9.894
	418	MFFNPEESELDLTYG	CMGC/CLK/CLK1	3.75	2.638
	418	MFFNPEESELDLTYG	Other/CDC7/CDC7	12.667	11.933
	418	MFFNPEESELDLTYG	Other/CDC7	12.667	11.933
	418	MFFNPEESELDLTYG	Other/PEK/HRI/EIF2AK1	5.333	4.658

	418	MFFNPEESELDLTYG	Other/PLK/PLK1/PLK1	3.791	3.783
	418	MFFNPEESELDLTYG	Other/ULK/ULK/ULK1	5.5	4.508
	423	EESELDLTYGNRYKD	CAMK/CAMKL/PASK/PASK	7.5	7.492
	423	EESELDLTYGNRYKD	Other/TTK/MPS1	7.333	7.039
	423	EESELDLTYGNRYKD	Other/TTK	5.375	5.303
	423	EESELDLTYGNRYKD	STE/STE20/FRAY	5.692	4.646
	423	EESELDLTYGNRYKD	STE/STE20/FRAY/STLK3	6	5.467
	424	ESELDLTYGNRYKDV	TK/Eph/EphA4	11	8
	424	ESELDLTYGNRYKDV	TK/Syk/SYK	7.11	6.754
	424	ESELDLTYGNRYKDV	TK/Tec/BMX	6.4	3.05
	424	ESELDLTYGNRYKDV	TK/VEGFR/FLT1	5.167	4.905
	428	DLTYGNRYKDVKLPD	TK/Eph/EphA4	9.25	8
	428	DLTYGNRYKDVKLPD	TK/Fer/FEs	9.091	7.262
	428	DLTYGNRYKDVKLPD	TK/Fer	8.474	7.78
	448	ILDVFCGSQMHIVRS	Atypical/PIKK/ATM/ATM	4.299	3.362
	448	ILDVFCGSQMHIVRS	Atypical/PIKK/ATM	4.256	4.149
	448	ILDVFCGSQMHIVRS	Atypical/PIKK/ATR	5.702	5.485
	448	ILDVFCGSQMHIVRS	CAMK/MAPKAPK/MNK/MNK2	4	4
	448	ILDVFCGSQMHIVRS	CMGC/DYRK/DYRK2/DYRK3	8.667	7.517
	448	ILDVFCGSQMHIVRS	Other/NEK/NEK2/NEK2	5.3	5.105
	448	ILDVFCGSQMHIVRS	Other/NEK/NEK2	5.3	5.105
	455	SQMHIVRSDELREAW	Atypical/Alpha/ChaK	6	3.588
	455	SQMHIVRSDELREAW	Atypical/Alpha/ChaK/TRPM7	12	6.629
	455	SQMHIVRSDELREAW	Atypical/Alpha	3.375	1.988
	455	SQMHIVRSDELREAW	CAMK/MAPKAPK/MNK/MNK2	4	4
	466	REAWRIFTPLLHHIE	AGC/PKA/PKACB	6.333	6.283
	466	REAWRIFTPLLHHIE	Atypical/PIKK/FRAP/TOR2	5.714	4.986
	466	REAWRIFTPLLHHIE	CMGC/CDK/CDK2/CDK2	13.846	12.788
	466	REAWRIFTPLLHHIE	CMGC/CDK	6.438	5.994
	466	REAWRIFTPLLHHIE	CMGC/DYRK/HIPK/HIPK2	2.276	1.841
	466	REAWRIFTPLLHHIE	CMGC/MAPK/nmo/NLK	3.812	3.509
	466	REAWRIFTPLLHHIE	CMGC/MAPK/nmo	3.938	3.141

	466	REAWRIFTPLLHHIE	STE/STE20/PAKA/PAK2	2.625	2.05
	482	EKARPIPYVYGSRGP	TK/Eph/EphB5	2.167	2.142
	482	EKARPIPYVYGSRGP	TK/PDGFR/PDGFRA	11.8	11.527
	484	ARPIPYVYGSRGPVE	TK/Eph/EphB2	4.929	4.611
	484	ARPIPYVYGSRGPVE	TK/PDGFR/FLT3	1.444	1.067
	486	PIPYVYGSRGPVEAD	CAMK/CAMKL/MARK/MARK1	7.444	7.433
	486	PIPYVYGSRGPVEAD	CAMK/MAPKAPK/MNK/MNK2	4	4
	486	PIPYVYGSRGPVEAD	Other/NEK/NEK11/NEK11	5.667	5.067
	486	PIPYVYGSRGPVEAD	Other/PLK/PLK2/PLK2	6.96	6.272
	486	PIPYVYGSRGPVEAD	Other/TLK	4.75	2.312
	503	MKRVGFQYEGTYKWV	TK/Csk/CTK	14	11.4
	503	MKRVGFQYEGTYKWV	TK/Eph	10.024	9.284
	503	MKRVGFQYEGTYKWV	TK/Syk	1.825	1.813
	503	MKRVGFQYEGTYKWV	TK/VEGFR/FLT1	5.333	4.905
	506	VGFQYEGTYKWVNPH	AGC/GRK/BARK/BARK1	12.857	10.916
	506	VGFQYEGTYKWVNPH	AGC/GRK/BARK	5.488	4.78
	506	VGFQYEGTYKWVNPH	AGC/PKC/PKCa/PRKCA	14.46	7.372
	506	VGFQYEGTYKWVNPH	AGC/PKC	3.194	1.416
	506	VGFQYEGTYKWVNPH	Atypical/PDHK/PDHK	9.15	5.455
	506	VGFQYEGTYKWVNPH	Other/TTK	5.625	5.303
	506	VGFQYEGTYKWVNPH	Other/WEE/Myt1/PKMYT1	5.667	5.583
	506	VGFQYEGTYKWVNPH	Other/WEE/Myt1	5.667	5.583
	506	VGFQYEGTYKWVNPH	Other/WEE	5.667	5.583
	506	VGFQYEGTYKWVNPH	STE	5.14	1.285
	506	VGFQYEGTYKWVNPH	STE	5.14	1.285
	506	VGFQYEGTYKWVNPH	STE/STE20/FRAY	5.077	4.646
	506	VGFQYEGTYKWVNPH	STE/STE20/FRAY/STK39	7.556	6.35
	506	VGFQYEGTYKWVNPH	STE/STE20/MST/MST1	48.355	9.234
	506	VGFQYEGTYKWVNPH	STE/STE20/MST	28.342	9.532
	506	VGFQYEGTYKWVNPH	TKL/IRAK/IRAK1	6	2.8

	506	VGFQYEGTYKWVNPH	TKL/IRAK	3	2.088	
	506	VGFQYEGTYKWVNPH	TKL/LISK	6	5.388	
	507	GFQYEGTYKWVNPHK	TK/DDR/DDR2	4.333	2.333	
	507	GFQYEGTYKWVNPHK	TK/DDR	4.333	2.333	
	507	GFQYEGTYKWVNPHK	TK/Eph	12.024	9.284	
	507	GFQYEGTYKWVNPHK	TK/FGFR/FGFR2	2.333	1.775	
	507	GFQYEGTYKWVNPHK	TK/Fer/FER	14.7	7.185	
	507	GFQYEGTYKWVNPHK	TK/PDGFR/CSF1R	8.867	2.947	
	507	GFQYEGTYKWVNPHK	TK/Src/SRM	1.545	1.527	
	507	GFQYEGTYKWVNPHK	TK/Src/SrcA	2.608	2.268	
	507	GFQYEGTYKWVNPHK	TK/Src/SrcB/LYN	7.937	5.358	
	507	GFQYEGTYKWVNPHK	TK/Src	1.643	1.63	
	507	GFQYEGTYKWVNPHK	TK/Tec/ITK	4.1	2.44	

Supplementary Figure 1: Predicted conservation profile of the ovine G6PD protein residues.