Egyptian Journal of Medical Human Genetics

META-ANALYSIS Open Access

D76V, L161R, and C117S are the most pathogenic amino acid substitutions with several dangerous consequences on leptin structure, function, and stability

Mohammed Baqur S. Al-Shuhaib

Abstract

Background: Leptin is a versatile hormone with a variety of functions, including regulation of food intake by inhibiting hunger. Any deleterious mutation in this protein can lead to serious consequences for the body. This study was conducted to identify the most deleterious non-synonymous single-nucleotide polymorphisms (nsSNPs) of human *LEP* gene and their impact on its encoded protein.

Methods: To predict the possible impact of nsSNPs on leptin, a total of 90 nsSNPs were retrieved from dbSNP and investigated using many in silico tools which specially designed to analyze nsSNPs' consequences on the protein structure, function, and stability.

Results: Three nsSNPs, namely D76V, L161R, and C117S, were found to be completely deleterious by all utilized nsSNPs prediction tools, thus affecting leptin protein structure, biological activity, and stability. Evolutionary information indicated L161R and C117S mutations to be located in extremely high conserved positions. Furthermore, several deleterious mechanisms controlled by both L161R and C117S mutations which alter several motifs in the secondary structure of leptin were detected. However, all D76V, L161R, and C117S mutations exhibited alteration in polar interactions in their representative positions. Further in-depth analyses proved several harmful structural effects of the three nsSNPs on leptin, which may lead to multiple intrinsic disorders in the altered protein forms.

Conclusions: This study provides the first comprehensive computation of the effect of the most damaging nsSNPs on leptin. The exploration of these missense mutations may present novel perspectives for various deleterious consequences originated from such amino acids substitutions. The dynamics of leptin performance, therefore, in many biological pathways, may be changed to create a variety of disorders, such as obesity and diabetes. These findings will help in detecting the most harmful variations needed to be screened for clinically diagnosed patients with leptin disorders.

Trial registration: ISRCTN73824458

Keywords: Consequences, Human, In silico, Non-synonymous, Obesity, Protein, SNPs

Correspondence: mohammed79@agre.uoqasim.edu.iq; baquralhilly_79@yahoo.com
Department of Animal Production, College of Agriculture, Al-Qasim Green University, 8-Al Qasim, Babil 51001, Iraq

Background

Leptin is the product of the *LEP* gene, a multifunctional hormone that participates in a variety of cellular activities, including controlling body weight, energy homeostasis, appetite, and reproduction. The mature form of leptin is composed of 16 kDa, which is secreted into the blood circulatory system by adipocytes and plays a versatile role in many metabolic pathways in the body [1]. Its deficiency can lead to profound diabetes, obesity, and infertility [2]. Leptin is encoded by the *LEP* gene, which is positioned in q32.1 locus within chromosome 7 and spans approximately 20 kb of DNA sequences. It consists of three exons separated by two introns. Actually, the first exon is noncoding and it is truncated in the mature blood circulating hormone, while the other two exons produce the fully mature 167 residues of blood circulating leptin (Chromosome 7–NC_000007.14). Leptin contains distinctive three-dimensional (3D) four-α-helix bundle folds of A-B-C-D tertiary structure [3]. This structure is arranged in a four sequentially similar antiparallel left-hand twisted α-helices bundle that is connected by two crossover links, alongside with one short loop [4]. In addition to the four main helices, one disulfide bond has been found to connect two cysteine residues (Cys117–Cys167) within the C and D helices respectively to form a crucial kink, which has been proven to be important for the leptin integrity, and biological activity [5]. Thus, any amino acid substitution that alters this highly organized 3D structure may have a series of damaging consequences on the final manifestation of the altered protein. On the other hand, it is well known that mutations can induce several effects on the corresponding proteins either by changing the expression of the affected proteins by substituting the transcription factors [6], interfering with the splicing [7], or by single amino acid substitutions (or nsSNPs) [8]. However, in the latter case, which is present within the coding portion, an alternative amino acid is incorporated in the protein chain and is known to be one of the main causes of the possible alterations in the leptin mode of action, which may lead to several undesired consequences. Accordingly, it is important to differentiate these consequences computationally [9]. Wet lab studies of amino acid substitutions intended to identify their consequences on their corresponding proteins are laborious and expensive. Contrastingly, the recent cumulative computational tools provide robust, rapid, low-cost, and comprehensive insights into the mechanisms of these missense SNPs [10]. Although some in silico analyses have focused on the missense mutations in leptin protein [11], no comprehensive study has published yet to predict the final consequences of the whole amino acid substitutions in this protein. Therefore, this study is designed to provide the first in silico-based prediction for the all missense mutations in the human *LEP* gene to identify the most deleterious SNPs in terms of protein structure, function, and biological interactions.

Methods

Meta-analysis was carried out according to PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines (http://www.prisma-statement.org/).

Web servers used in computational prediction

The current computational study for the structural and functional effects of the observed nsSNPs was based on several in silico websites and servers, including dbSNP (https://www.ncbi.nlm.nih.gov/projects/SNP/), ensemble genome browser 95 (https://asia.ensembl.org), SIFT (Sorting Intolerant from Tolerant SNPs) (http://sift.bii.a-star.edu.sg/www/SIFT_seq_submit2.html), PolyPhen-2 (Polymorphism Phenotyping v2) (http://genetics.bwh.harvard.edu/pph2/), REVEL (Rare Exome Variant Ensemble Learner) (https://asia.ensembl.org/index.html), MetalR (https://asia.ensembl.org/index.html), PROVEAN (Protein Variation Effect Analyzer) (http://provean.jcvi.org/index.php), PANTHER (Protein ANalysis THrough Evolutionary Relationships) (http://www.pantherdb.org/tools/csnp), SNAP2 (Screening for Non-Acceptable Polymorphisms 2) (https://www.predictprotein.org), SNPs&GO (https://snps-and-go.biocomp.unibo.it/snps-and-go/), PhD-SNP (Prediction of Deleterious Single Nucleotide Polymorphism) (http://snps.biofold.org/phd-snp/phd-snp.html), CUPSAT (Cologne University Protein Stability Analysis Tool) (http://cupsat.tu-bs.de/), ConSurf (http://consurf.tau.ac.il/2016/), PolyView-2D (http://polyview.cchmc.org/), MutPred (http://mutpred1.mutdb.org/), UNIPROT (http://www.uniprot.org), SWISS-MODEL (https://swissmodel.expasy.org/assess/), RaptorX (http://raptorx.uchicago.edu/), I-TASSER (iterative threading assembly refinement algorithm) (https://beta.swissmodel.expasy.org/qmean/), Phyre2 (Protein Homology/analogY Recognition Engine) (http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index), VERIFY 3D (http://servicesn.mbi.ucla.edu/Verify3D/), PROCHECK (https://www.ebi.ac.uk/thornton-srv/software/PROCHECK/), PyMol-v1, 7.0.1 (www.shrodinger.com), TM-align (https://zhanglab.ccmb.med.umich.edu/TM-align/), HOPE (http://www.cmbi.ru.nl/hope/method/), and String 10 (https://string-db.org/).

Retrieval of SNPs data and protein sequence

The dbSNP database was initially used to retrieve the nsSNPs of the *LEP* gene analyzed in the present study. Subsequently, ensemble genome browser 95 was used to further filter out the retrieved SNPs. After categorizing the retrieved SNPs, all nsSNPs were considered to prioritize their structural, functional, and stability effects on leptin by utilizing different missense state-of-

the-art in silico prediction tools. Leptin amino acid sequences with national center of biotechnology information (NCBI) accession number XP_005250397.1 was the input protein FASTA sequences.

Finding deleterious nsSNPs by SIFT

The structural consequences of all the retrieved nsSNPs in their corresponding positions in the human leptin protein were analyzed using SIFT program. Substitutions at each position with less than a tolerance index of 0.05 were predicted as "intolerant" or "deleterious", while those greater than or equal to 0.05 as "tolerated" [12].

Evaluating the functional and structural impact of nsSNPs by Polyphen-2

Polyphen-2 was utilized to analyze the possible effect of an amino acid nsSNP on structure, as well as the function of the analyzed protein by means of multiple sequence alignment [13]. Three common scores were obtained by this software; "probably damaging", "possibly damaging", and "possibly benign" based on the scores that are ranged from "0" to "1" respectively.

Pathogenicity prediction using REVEL tool

The potential pathogenicity of the amino acid substitutions was assessed by REVEL, a recently developed ensemble software for discriminating between neutral and pathogenic amino acid substitutions on the basis of several in silico tools. Spearman rank correlation coefficient (R) values of > 0.6 indicate high pathogenicity, while $0.4 < R < 0.6$ indicates moderate pathogenicity, and $R < 0.4$ indicate low pathogenicity of the missense mutation [14].

Pathogenicity prediction using MetalR tool

Another validation of the potential pathogenicity of leptin missense mutations came from MetalR software, a tool for predicting the pathogenicity of nucleotide mutation through a logistic regression based ensemble method. In MetalR, the amino acid mutation is classified as 'tolerated' or 'damaging'; a score between 0 and 1 [15]. The pathogenicity predictions of leptin missense mutations for both REVEL and MetalR tools were retrieved from ensemble browser 95 genome.

Validating the deleterious nsSNPs through PROVEAN

The biological consequences of the observed mutations were validated using PROVEAN, which is a prediction machine that separates between the neutral and deleterious amino acids, by relying on a threshold of −2.5, the substitution predicted as deleterious when it scores is less than ≤ −2.5 [16].

Finding the biological impact of nsSNPs by PANTHER

Validation of SIFT results was performed by PANTHER tool, a bio-computational tool that estimates the evolutionary probability of each nsSNP to have a biological impact on the evaluated protein of interest [17]. The expected Panther scores of each nsSNP are "probably damaging" (when time > 450 my), "possibly damaging" (when 450 my > time > 200 my), and "probably benign" (when time < 450 my).

Prediction of deleterious nsSNPs by SNAP2

SNAP2 is an amino acid substitution prediction tool that provides more confirmative data on the functional consequences of the missense mutation in its corresponding position in the whole protein. SNAP2 scores range from a damaging "effect" to non-damaging "neutral" scores when it gives > zero and < zero scores respectively [18].

Predicting amino acid mutations association with diseases using SNPs&Go

Predicting Human Disease-Related Mutations in leptin was performed using SNPs&GO, a server for the prediction of single point protein mutations likely to be involved in the development and progression of diseases. The particular mutation is disease-causing when it is scored greater than 0.5 (> 0.5) [19].

Predicting disease-causing amino acid substitutions using PhD SNP

The potential pathogenicity of the missense mutations was also validated using PhD SNP, a support vector machine-based detector of human deleterious SNP. It predicts whether the given missense mutation leads to a disease development according to the reliability index score. The same SNPs&GO scores were based to assess the pathogenicity of the mutation [20].

Analysis of leptin evolutionary conservative regions using ConSurf

ConSurf server was a computational machine utilized to assess the evolutionary status of the amino acids in a particular protein. In ConSurf tool, several algorithm-based alignments were made to predict whether the analyzed amino acid residue was variable, neutral, or conserved within the analyzed protein to give rise to nine grades of color spectrum, each one refers to the approximate conservation status of the analyzed residues [21]. Subsequently, further details of the observed extremely harmful nsSNPs were provided by generating a 2D structure of leptin using PolyView-2D tool.

Predicting functional properties of the deleterious amino acids substitutions by Mutpred

The most deleterious amino acid substitutions consequences were further assessed as deleterious or neutral using the MutPred, a web server application to predict the mechanism used by a missense mutation to interfere with protein biological activity. The amino acid mutation with a score greater than 0.5 is predicted as deleterious [22].

3D Modeling and structural analysis of leptin

The UniProtKB/Swiss-Prot entry number of this protein is Q4TVR7 and the UniProt accession number of human leptin, P41159, was used as an input protein data. No matching protein data bank (PDB) entries were found in this server. Therefore, the 3D structure of human leptin was built by comparing three available 3D modelling servers, including RaptorX, a web portal for 3D generating tertiary structure, solvent accessibility, contact map, and binding sites of the protein FASTA (text-based format for representing either nucleotide sequences or amino acid sequences) sequence [23], I-TASSER [24], and Phyre2, [25]. The best stereochemical properties of each generated 3D models were validated using PROCHECK online server [26]. After choosing the best model, the observed amino acid substitutions and the possible alterations in polar contacts were analyzed by PyMol-v1, 7.0.1.

Predicting amino acid mutation effect on protein stability through CUPSAT

To get a better view to the stability of the mutant human leptin, amino acids substitutions were analyzed using CUPSAT, an automatic web server for the assess the protein stability changes upon point mutations [27]. The RaptorX—built input PDB file of the human leptin was used as an input file and computed in terms of free energy change (DDG) value (kcal/mol).

Superimposition of leptin wild type with its deleterious nsSNPs using TM-align

TM-align was used to compare wild-type protein structure with its mutant counterparts [28]. This software measure template modeling-score (TM-score) and root-mean-square deviation (RMSD) along with superposition of the 3D built structures. TM-score gives the values in 0 and 1, where 1 indicates a perfect match between two structures, while higher RMSD values indicate a greater variation between wild-type and mutant structures.

In-depth structural analysis and interactions of deleterious nsSNPs using HOPE and String 10 servers

The virtual observations for analyzing the structural effects of the most harmful nsSNPs were obtained using HOPE tool, an automatic mutant analysis server which builds an animated report that is easy to understand with in-depth analyses of the targeted amino acid substitution on the targeted protein [29]. Subsequently, the possible protein-protein interaction analysis was conducted by String 10 web server [30].

Results and discussion

In this study, a series of in silico prediction analyses were utilized to find out most deleterious amino acid substitutions in the coding region of the *LEP* gene from other substitutions obtained from the dbSNP database by employing multiple computational tools and then observing their effect on structure, function, stability, as well as the evolutionary conservation scores with regard to their corresponding amino acid residues in leptin.

Retrieval of nsSNPs and screening the most deleterious mutations

A total of 3881 SNPs (including 1 mutation of near 5′-untranslated region (UTR), 14 mutations of 5′-UTR, 76 synonymous mutations, 90 nonsynonymous mutations, 5 frameshift mutations, 1 inframe deletion, 2 mutations of splice acceptor, 1 splice donor mutations, 3 stop gained mutations, 2939 intronic mutations, 15 splice region mutations, and 734 3′-UTR mutations) that were discovered in the human *LEP* gene, only nsSNPs were screened in this study (Table 1).

Screening of the retrieved amino acids to screen the most deleterious mutations

After retrieval of all amino acid substitutions, ten different computational tools were used for the structural and functional annotation of these nsSNPs, including SIFT [12], PolyPhen-2 [13], REVEL [14], MetalR [15], PROVEAN [16], PANTHER [17], SNAP2 [18], SNPs&Go [19], PhD-SNP [20], and CUPSAT [27] respectively, which were used to evaluate the potential consequences of their deleterious effects on protein structure, function, and stability. By comparing the computation of these methods, the degree of the damaging consequences of each individual missense mutation was verified. Thus, the structural and biological outcomes of each amino acid substitution were assessed using the cumulative prediction of these tools. However, sometimes, variable results were observed for the same mutation. This could be explained by the differences in the algorithm based on each tool to resolve the outcomes. However, the extensive cumulative predictions gave clear confirmations with regard to the extent of the deleterious or neutral effects of each amino acid mutation on the leptin. These confirmations came from the ability of the utilized nsSNPs prediction tools to give compatible results in the collective effects of the currently analyzed 90 missense mutations (Additional file 1: Table S1). Out of the 90

Table 1 Distribution of the retrieved SNPs of the bovine *LEP* gene. The bolded SNPs were the nonsynonymous SNPs that are selected for the present comprehensive study

No.	Type of SNP	No. of SNPs
1	Start retained variant	1
2	5′ UTR variants	14
3	Synonymous variants	76
4	**Missense (non-synonymous)variants**	**90**
5	Frameshift variants	5
6	Inframe deletion	1
7	Splice acceptor variant	2
8	Splice donor variant	1
9	Stop gained	3
10	Intronic variants	2939
11	Splice region variants	15
12	3′ UTR variants	734
Total		3881

amino acid substitutions, only three were found to be deleterious by ten different computational tools that were involved for the assessment of whether the analyzed amino acid mutation has a harmful or neutral effect on the protein. The results showed that only three amino acid substitutions, D76V, C117S, and L161R, were found to be subject to absolutely confirmed deleterious, damaging, or pathogenic effects by all the utilized prediction tools (Table 2).

Thus, all the prediction tools that were concerned with predicting structure, function, and stability were in line with each other with regard to the confirmed deleterious effects of the D76V, C117S, and L161R (Fig. 1), while any other nsSNP that was not exhibited deleterious consequences by all the mentioned in silico tools were omitted from further analyses.

The ranges of the deleterious consequences of D76V, C117S, and L161R mutants were further classified according to the mechanism of each mutation to effect on the biological structure and function of the altered leptin. ConSurf tool was provided evolutionary conservation analysis for these three harmful substitutions. It showed that both C117S and L161R mutations were located in two highly conserved regions within the primary evolutionary structure of leptin (Fig. 2a). Whereas, D76V mutation was located in a low conserved position, possibly indicating a less serious evolutionary damaging consequences caused by this mutation in comparison with C117S and L161R mutations. The variable evolutionary positioning of these loci suggested multiple dynamic

Table 2 List of the completely deleterious variants observed in the human *LEP* gene by extensive prediction approach

Tool	Parameter	Amino acid substitution (SNP ID)		
		D76V (rs1332916395)	C117S (rs1051206328)	L161R (rs771956117)
SIFT	Score	0	0	0
	Prediction	Deleterious	Deleterious	Deleterious
PolyPhen2	Score	1.000	1.000	1.000
	Prediction	Probably damaging	Probably damaging	Probably damaging
REVEL	Score	0.648	0.777	0.771
	Prediction	Deleterious	Deleterious	Deleterious
MetaLR	Score	0.569	0.839	0.509
	Prediction	Deleterious	Deleterious	Deleterious
PROVEAN	Score	-5.78	-8.46	-4.03
	Prediction	Deleterious	Deleterious	Deleterious
PANTHER	Score	220	220	220
	Prediction	Probably damaging	Probably damaging	Probably damaging
SNAP2	Score	70	72	81
	Prediction	Effect	Effect	Effect
SNPs&Go	Score	7	7	1
	Prediction	Disease	Disease	Disease
PhD SNP	Score	5	5	3
	Prediction	Disease	Disease	Disease
CUPSAT	Score (kcal/mol)	-1.5	-1.71	-2.91
	Prediction	Decrease stability	Decrease stability	Decrease stability

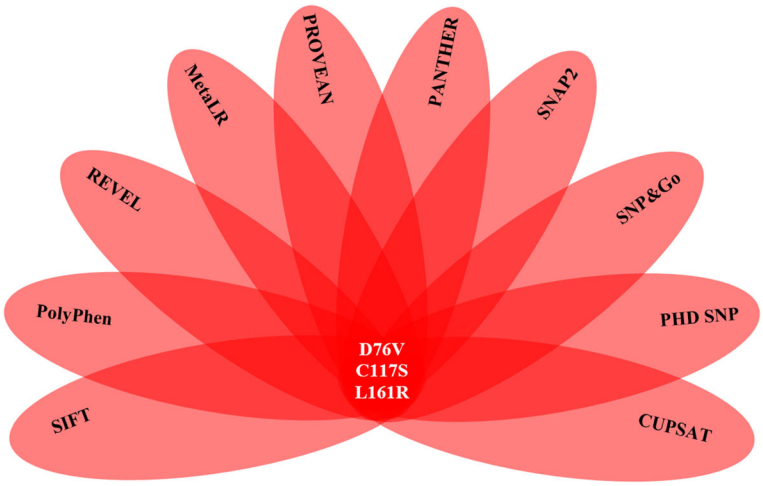

Fig. 1 Ten different computational tools utilized to predict the most deleterious effect of amino acid substitutions on leptin structure, function, and stability. These tools cumulatively showed that the most deleterious substitutions are D76V, C117S, and L161R

mechanisms played by these mutations in the alteration of the affected leptin. However, the ConSurf prediction of high critical roles of both C117S and L161R was confirmed by MutPred tool by predicting two crucial alterations of C117S and L161R on the mutant proteins. Although MutPred tool indicated no participation of D76V in inducing any possible alteration, it did find that C117S mutation induced altered metal binding and L161R induced remarkable disordered structure, reduced stability, and transmembrane localization (Table 3).

The possible explanation for these critical roles came from the critical positioning of C117S and L161R in the secondary structure of leptin, which showed that C117S is located in the Cys117-Cys167 disulfide bond, and, therefore, perhaps represents the most drastic nsSNP in comparison with L161R and D76V, respectively (Fig. 2b). In agreement with the previously mentioned bioinformatics tools, human leptin was found to have three functionally important receptor binding sites on the four-helix leptin structure [31]. The leptin receptor binding site-I is located in the C-terminus of helix D. It is a 50 amino acids long chain, which positioned within 117–167 residues in leptin. However, the helix D exhibits a characterized structure, which may permit leptin to bind specifically to a receptor through enhancing the activity of *N*-terminal through the Cys117–Cys167 disulfide bond [32]. Therefore, C117S mutation induced a deleterious modification in the binding site-I that specified for leptin receptor [33]. Thus, the current findings predicted that the mutation in this locus usually induces damaging consequences on leptin structure, biological activity, and stability through manipulating the referred Cys117-Cys167 disulfide bond.

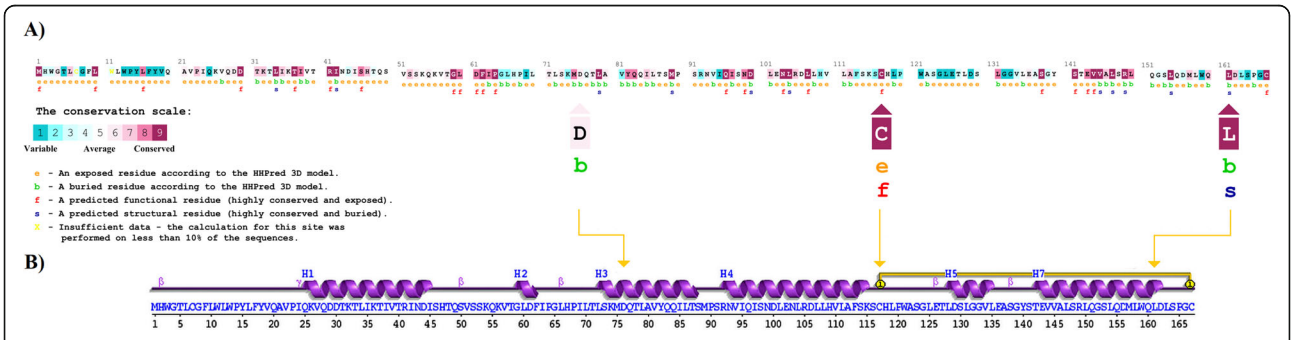

Fig. 2 The positioning of the most deleterious amino acid substitutions identified in leptin. **a**) Evolutionary conservation pattern of amino acid residues within the primary sequence of leptin. Color intensity increases with degree of conservation. The amino acids are colored based on their conservation grades and conservation levels. The amino acid positions of the most three deleterious predicted SNPs are maximized below the protein sequence. Both C117 and L161 residues showed higher conserved positions compared with D76 residue. **b**) Secondary structure prediction of leptin. Both D76 and L161 amino acid residues are positioned in the alpha-helix region, while the C117 amino acid residue is suited in the highly critical cysteine-cysteine junction. The primary and secondary structures are generated by ConSurf and PolyView-2D tools, respectively

Table 3 MutPred prediction for the most deleterious missense variants observed in *LEP* gene

Variant	MutPred2 score	Remarks	Affected PROSITE and ELM Motifs
D76V	0.436	–	–
C117S	0.735	-	ELME000063
	Molecular mechanisms with *P* values ≤ 0.05	Probability	*P* value
	Altered metal binding	0.39	6.9e–03
L161R	0.720	–	ELME000106, ELME000136, ELME000159
	Molecular mechanisms with *P* values ≤ 0.05	Probability	*P* value
	Gain of intrinsic disorder	0.36	0.02
	Altered stability	0.21	0.01
	Altered transmembrane protein	0.14	0.02

Homology modeling validation of leptin

To determine the extent to which the three deleterious nsSNPs can alter the wild-type structure of the leptin protein, a 3D structure of leptin was built. The efficiency of the three modeling servers was compared to generate the most appropriate tertiary structure, including RaptorX, PhyRe2, and I-TASSER. It was found that the RaptorX tool was found to have built the best 3D models. Furthermore, Ramachandran plot statistics of the generated 3D structure of leptin by RaptorX revealed that the standard deviation (omega angle degree), bad contacts per 100 residues, zeta angle degree, hydrogen bonds energies, and overall G-factor were scored values of 3, 2.4, 1.4, 0.6, and 0.1 respectively which were under the control and accepted limit in comparison to both PhyRe2 and I-TASSER-generated protein structures, respectively (Fig. 3).

In the RaptorX generated model, 167 (100%) of amino acid residues were modeled, in which only 14 (8%) of positions were predicted as disordered. This generated model has 61% helix, 1% strand, and 37% coil. The result of

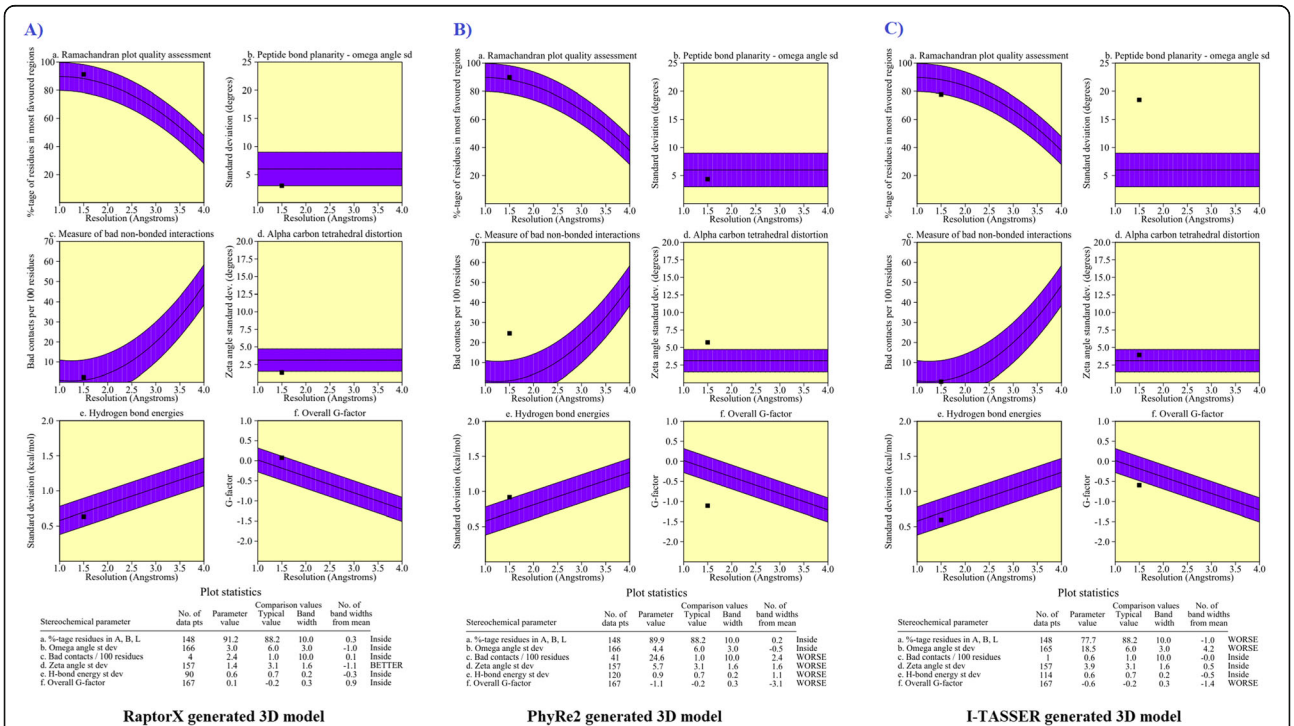

Fig. 3 Stereochemical comparison among three common 3D structure modeling servers in terms of generating the best tertiary structure of leptin. **a**) RaptorX generated leptin. **b**) PhyRe2 generated leptin. **c**) I-TASSER generated leptin. All parameters revealed from RaptorX built structures showed accepted limit values and more qualified parameters compared with PhyRe2 and I-TASSER built structures, respectively. The 3D structure is validated by side chain parameters of PROCHECK tool

solvent accessibility of secondary structure of leptin was 47% intermediate, 25% medium, and 26% buried. Ramachandran plot was used to validate the leptin protein model obtained from the RaptorX. Out of 167 amino acids, 135 (91.2%) were in the most favored region, 13 (8.8%) in additional allowed regions, while no residues were found in the disallowed regions. The number of non-glycine and non-proline residues was found to be 148 (100%). End residues (excl. Gly and Pro) were only 2, glycine residues (shown in triangles) were 10, and the number of proline residues was 7 (Additional file 2: Fig. S2).

Structural analysis of the most deleterious mutations

Using RaptorX generated models, the superimposed structure of each one of D76V, C117S, and L161R deleterious amino acid substitutions was performed in the native 3D structure. The current analyses were extended further by calculating the TM-scores and RMSD values for each mutant model. The TM-score is used to evaluate the topological similarity between wild-type and mutant models, while the RMSD measures the average distance between α-carbon backbones of wild-type and mutant models [28]. The greater the RMSD value the greater is the deviation of mutant structure from that of the wild type which in turn changes their functional activity [34]. The mutant model for L161R showed maximum RMSD value which followed by those of D76V and C117S respectively, indicating more alteration forces of L161R compared with the other two D76V and C117S mutation forms (Table 4).

Table 4 TM-align predictions for the most three deleterious nsSNPs in leptin

NsSNP ID	Amino acid change	Aligned length	TM-score	RMSD (Å)
rs1332916395	D76V	153	0.82942	0.948
rs1051206328	C117S	166	0.93578	0.940
rs771956117	L161R	152	0.83738	0.967

The structural alterations between both wild-type and mutated leptin were visualized using the same 3D PDB models to visualize the positions of these deleterious mutations of the 3D models before and after mutations. However, a clear image of the wild-type

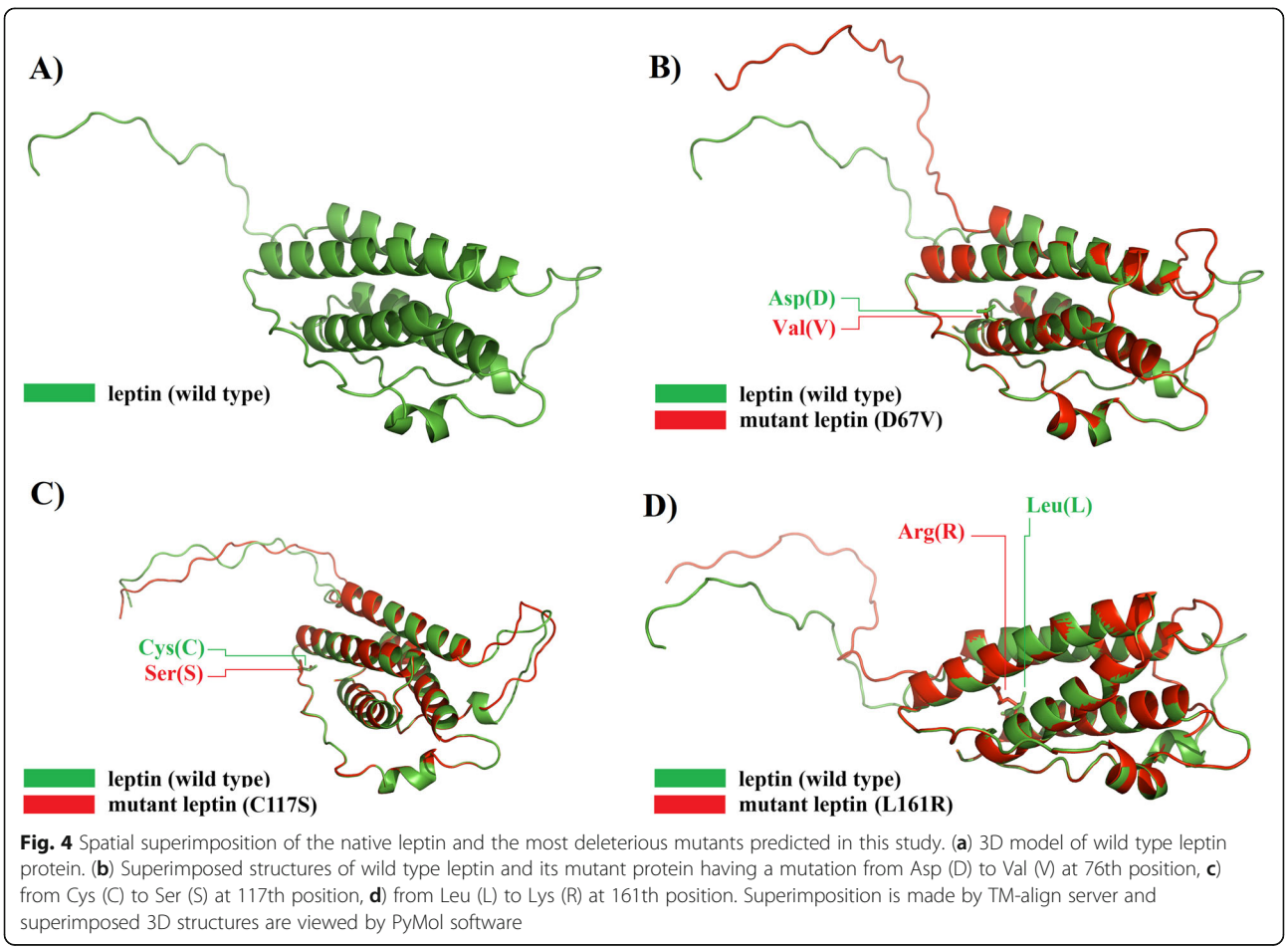

Fig. 4 Spatial superimposition of the native leptin and the most deleterious mutants predicted in this study. (**a**) 3D model of wild type leptin protein. (**b**) Superimposed structures of wild type leptin and its mutant protein having a mutation from Asp (D) to Val (V) at 76th position, **c**) from Cys (C) to Ser (S) at 117th position, **d**) from Leu (L) to Lys (R) at 161th position. Superimposition is made by TM-align server and superimposed 3D structures are viewed by PyMol software

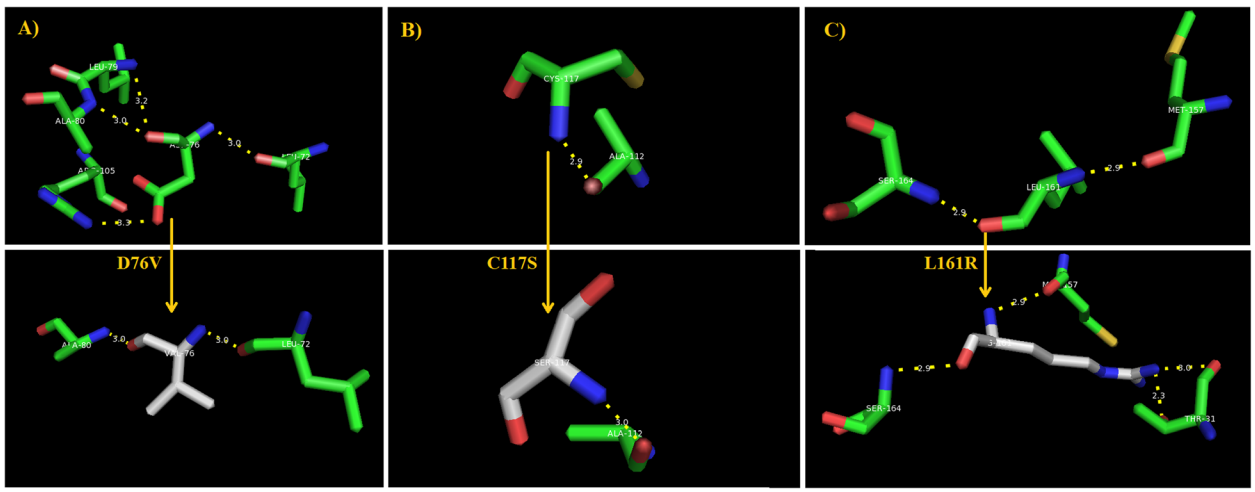

Fig. 5 Polar interactions of 3D structures of leptin before mutation and after its being mutated with D76V, C117S, and L161R, in **a**), **b**), and **c**), respectively. Noticeable alterations in the binding capacity between leptin wild-type residue and both D76V and L161R mutants' residues are found. The C117S mutant residue showed the same binding capacity compared with the wild type residue and the only alteration is found in the distances of this binding. Figures are generated by the measurement tool of PyMol software

and mutant leptin structures were provided by superimposing their 3D structures using RaptorX generated models, which explicitly showed the structural alterations upon mutations (Fig. 4).

The patterns of the polar interaction of the most deleterious D76V, C117S, and L161R mutations were assessed by PyMol to explain their expected roles in the conversion of the nature of the native amino acid binding with its surrounding residues in leptin [8]. However, the nature of the polar interactions of D76V, C117S, and L161R amino acid residues in the native protein and its mutant counterpart were investigated by PyMol to unravel its possible role in the mutant leptin. It was found that D67 in the native protein had four polar interactions with Leu72, Leu79, Ala80, and Arg105 of the following lengths: 3.0 Å, 3.2 Å, 3.0 Å, and 3.3 Å lengths, respectively. Contrarily, the altered amino acid V76 had only two polar interactions with Leu72 and Ala80 with the same measurements (Fig. 5a). With regard to C117S, only one polar interaction was seen in both cases with Ala112, but in a different measurement before and after mutation (Fig. 5b). The polar interactions of the native L161 were observed with Met157 and Ser164 of even 2.9 Å distance with both residues, while the altered amino acid R161 had two extra interactions with Thr31 residue (Fig. 5c).

The highly deleterious D76V, C117S, and L161R mutants showed remarkable modifications in the polar binding pattern upon mutation, indicating the involvement of polar interactions of these mutations in the altered protein metabolic activity. Noticeable alterations in the polar binding patterns were observed in terms of inducing changes in numbers and distances of polar interactions due to the observed D76V, C117S, and L161R mutations. However, the present finding showed the polar interactions of these deleterious amino acid substitutions exhibited several changes upon mutations, indicating a possible role for these polar interactions in the interaction of the *LEP* gene with other genetic platforms.

Further details with respect to the in-depth structural effects of these harmful nsSNPs were provided from HOPE tool [29]. HOPE prediction tool has reconfirmed the PyMol's predicted polar interaction alterations by revealing a difference in charge between the wild-type and mutant amino acid in the mutant C76V and indicated that the charge of the buried wild-type residue is lost by this mutation. The wild-type, C76, and mutant amino acid, V76, differ in size in such a way that the mutant residue is smaller than the wild-type residue (Fig. 6a). Moreover, HOPE tools predicted that the mutation C76V caused an empty space in the core of the protein. The hydrophobicity of the wild-type and mutant residue will differ, and the mutation will cause loss of hydrogen bonds in the core of the protein; as a result, this will disturb correct folding. With regard to C117S, HOPE tool found that the hydrophobicity of the wild-type and mutant residue differs and the substitution of Cys with Ser in the 117th position may cause loss of hydrophobic

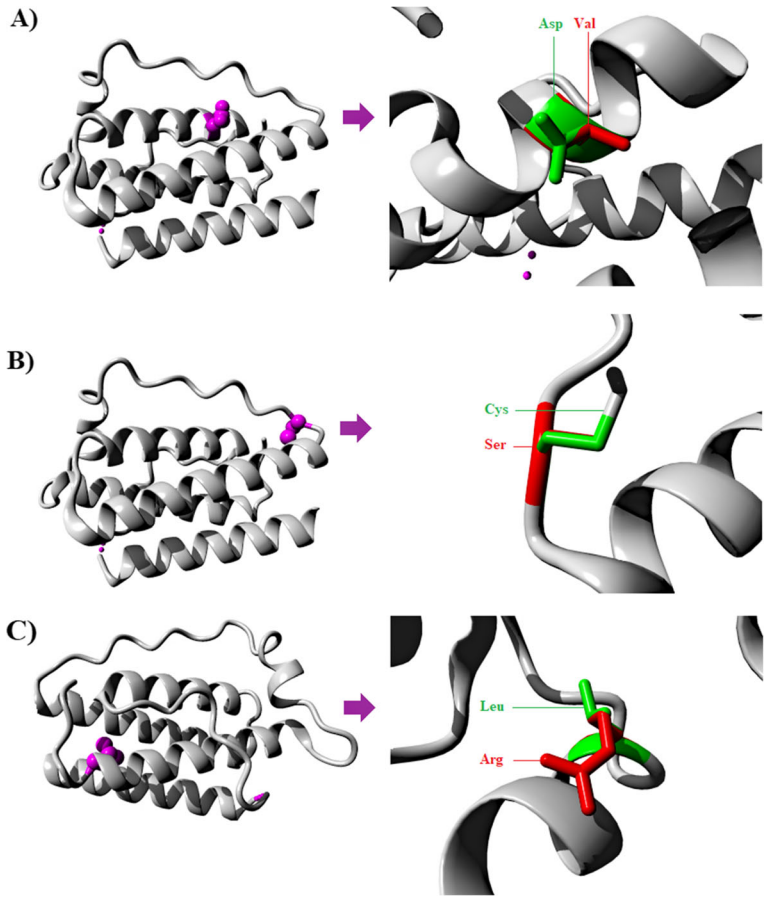

Fig. 6 The most deleterious amino acid substitutions predicted in leptin 3D structures. **a**) The wild type D and the mutant V residues at 78th position (D78V). **b**) The wild type C and the mutant S residues at 117th position (C117S). **c**) The wild type L and the mutant R residues at 161th position (L161R). The protein is shown in ribbon-presentation and colored grey. The wild type amino acid residues affected by mutations are shown in green, while their mutant counterparts are shown in red. Structures are generated by HOPE server

interactions with other molecules on the surface of the protein (Fig. 6b). In the case of L161R, HOPE detected a difference in charge between the wild-type and mutant amino acid. This difference came from the introduction of a charge in a buried residue that can lead to protein-folding problems. Furthermore, the wild-type and mutant amino acids differ in size since the mutant residue is bigger than the wild-type residue. As the wild-type residue was buried in the core of the protein, the mutant residue is bigger and probably will not fit (Fig. 6c). Therefore, the L161R mutation may lead to loss of hydrophobic interactions in the core of the protein. However, all three deleterious nsSNPs were found to be located in domains important for binding of other molecules and in contact with the residues in another domain. Thus, it is possible that the mutation in these residues can disturb these contacts. The mutation may affect the function of the protein, thereby disturbing the signal transfer from the binding domain to the activity domain.

These observed structural changes of mutant proteins indicate potential alterations in the binding affinity of mutant structures of leptin proteins with their receptors and substrates which ultimately lead to aberrant metabolism. Our results from String 10 server show that the *LEP* gene interacts with a variety of genes, mainly participating in the control of several cell cycle events (Fig. 7). Therefore, the currently highlighted deleterious nsSNPs may alter cell cycle controls by disrupting these genetic interactions, where leptin participates, with a particular emphasis on food intake regulation. However, disrupted leptin is incapable of undergoing its scheduled role of inhibiting hunger. This observation requires further attention to evaluate the percentage of such deleterious nsSNPs in patients with obesity and diabetes.

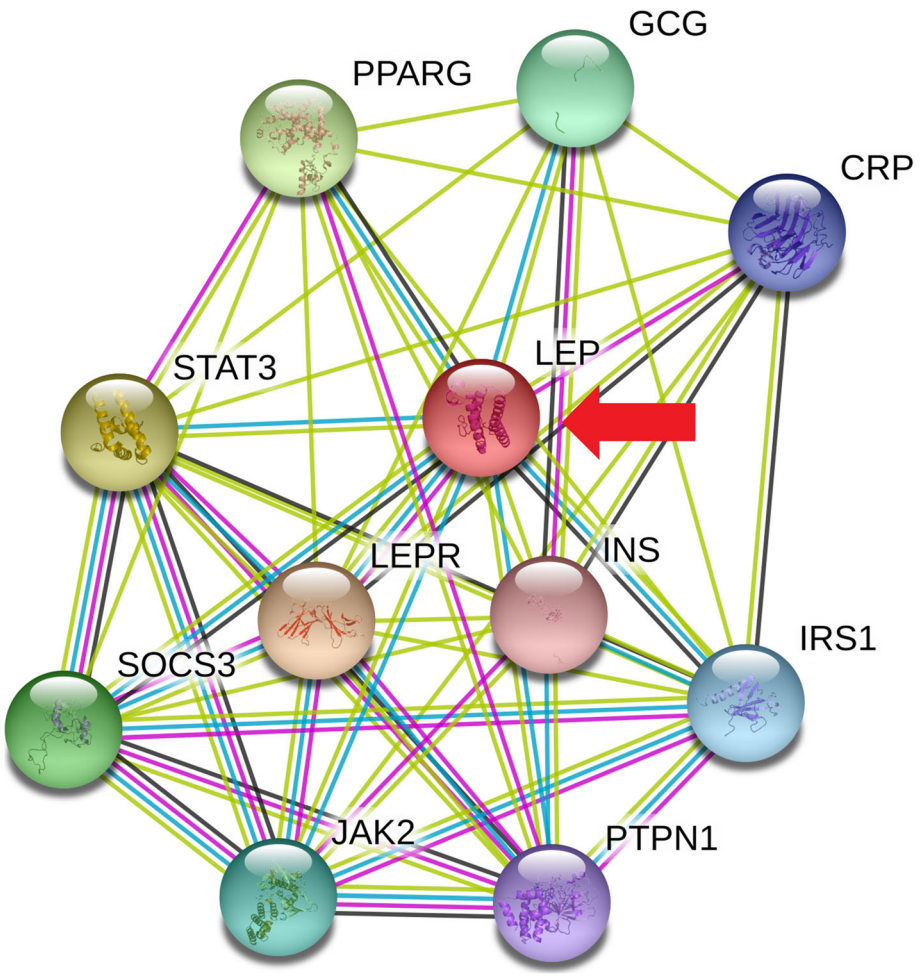

Fig. 7 Functional association network of the leptin with other important proteins in the cellular metabolic system. All known and predicted protein-protein interaction network are shown between leptin and other related proteins through variable extending threads. The red arrow refers to the targeted leptin protein. This network is generated by String 10 server

Conclusion

This study provides a decisive outcome concluding the accessible SNPs information by recognizing the three entirely harmful nsSNPs, namely rs1332916395 (D76V), rs1051206328 (C117S), and rs771956117 (L161R). Future studies should consider these nsSNPs as the main target mutations in various diseases involving *LEP* gene malfunctions.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s43042-019-0033-2.

Additional file 1: Table S1. List of all missense variants observed in the human *LEP* gene. The variants arranged according to the seriousness of their cumulative deleterious consequences. The green color refers to the neutral/stable effect of variant, while the red color refers to its deleterious/nonstable effect.

Additional file 2: Figure S1. The Ramachandran plot revealed that the phi/psi angles of 91.2% of the residues fell in the most favoured regions, 8.8% of the residues were in additional allowed regions, no residues were found in disallowed regions.

Abbreviations

CUPSAT: Cologne University Protein Stability Analysis TooIl-TASSERIterative Threading Assembly Refinement AlgorithmNCBINational Center of Biotechnology InformationnsSNPNon-synonymous Single Nucleotide PolymorphismPANTHERProtein ANalysis THrough Evolutionary RelationshipsPDBProtein Data BankPhD-SNPPrediction of Deleterious Single Nucleotide PolymorphismPhyRe2Protein Homology/analogY Recognition EnginePolyPhen-2Polymorphism Phenotyping v2PRISMAPreferred Reporting Items for Systematic Reviews and Meta-AnalysesPROVEANProtein Variation Effect AnalyzerREVELRare Exome Variant Ensemble LearnerRMSDRoot-mean-square-deviationSIFTSorting Intolerant from Tolerant SNPsSNAP2Screening for Non-Acceptable Polymorphisms 2TM-scoreTemplate Modeling-scoreUTRUntranslated Region

Acknowledgements

None.

Authors' contributions
The current manuscript was entirely performed and wrote by only one author and no other co-workers were involved. The author read and approved the final manuscript.

Funding
The present work was not funded.

Availability of data and materials
The current data were retrieved from dbSNP and ensemble genome browser 96.

Ethics approval and consent to participate
N/A

Consent for publication
N/A

Competing interests
The author declares that he has no competing interests.

Received: 26 June 2019 Accepted: 9 October 2019
Published online: 13 December 2019

References

1. Iserentant H, Peelman F, Defeau D, Vandekerckhove J, Zabeau L, Tavernier J (2008) Mapping of the interface between leptin and the leptin receptor CRH2 domain. J Cell Sci 118:2519–2527
2. Zhang F, Chen Y, Heiman M (2005) Dimarchi R (2005) Leptin: structure, function and biology. Vitam Horm 71:345–372
3. Londraville RL, Prokop JW, Duff RJ, Liu Q, Tuttle M (2017) On the molecular evolution of leptin, Leptin Receptor, and Endospanin. Front Endocrinol 8:58
4. Gutierrez DA, Puglisi MJ, Hasty AH (2009) Impact of increased adipose tissue mass on inflammation, insulin resistance, and dyslipidemia. Curr. Diabetes Rep 9:26–32
5. Haglund E, Sułkowska JI, He Z, Feng G-S, Jennings PA, Onuchic JN (2012) The unique cysteine knot regulates the pleotropic hormone leptin. PLoS ONE 7:e45654
6. Liao PY, Lee KH (2010) From SNPs to functional polymorphism: The insight into biotechnology applications. Biochem Eng J 49:149–158
7. Cartegni L, Chew SL, Krainer AR (2002) Listening to silence and understanding nonsense: exonic mutations that affect splicing. Nat Rev Genet 3(4):285–298
8. Al-Shuhaib MBS, Al-Kafajy FR, Badi MA, AbdulAzeez S, Marimuthu et al. Highly deleterious variations in *COX1*, *CYTB*, *SCG5*, *FK2*, *PRL* and *PGF* genes are the potential adaptation of the immigrated African ostrich population. Comput Biol Med 2018a;100:17–26.
9. Al-Shuhaib MBS, Al-Lamy SMA, Al-Tayy HMA, Al-Thuwaini TM, Radhi AH (2018b) Single Nucleotide Polymorphism (SNP) of leptin gene in holstein cattle **การแปรผันของลาดับดีเอ็นเอชนิดหนึ่ง (สนิป, SNP) ของยีนเลปตินในโคนมโฮลสไตน**. Thai J Vet Med 48(2):187–201
10. Abdulazeez S, Sultana S, Almandil NB, Almohazey D, Bency BJ, Borgio FG (2019) The rs61742690 (S783N) single nucleotide polymorphism is a suitable target for disrupting *BCL11A*-mediated foetal-to-adult globin switching. PLoS ONE 14(2):e0212492
11. Saranya GM, Prabhu F, Pathy MR (2018) Impression of missense single nucleotide polymorphisms of leptin gene on the early onset of obesity related infertility in female. Res J Biotechnol 13(9):35–47
12. Pauline CN, Steven H (2003) SIFT: predicting amino acid changes that affect protein function. Nucl Acids Res 31:3812–3814
13. Adzhubei I, Jordan DM, Sunyaev SR. Predicting Functional Effect of Human Missense Mutations Using PolyPhen-2. Curr Protoc Hum Genet 2013; Chapter 7;Unit 7.20.
14. Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S (2016) REVEL: an Ensemble method for predicting the pathogenicity of rare missense variants. Am J Hum Genet 99:877–885
15. Jagadeesh KA, Wenger AM, Berger MJ, Guturu H, Stenson PD, Cooper DN (2016) M-CAP eliminates a majority of variants of uncertain significance in clinical exomes at high sensitivity. Nat Genet 48(12):1581–1586
16. Choi Y, Sims GE, Murphy S, Miller JR, Chan AP (2012) Predicting the functional effect of amino acid substitutions and indels. PLOS ONE 7:e46688
17. Tang H, Thomas PD (2016) PANTHER-PSEP: predicting disease-causing genetic variants using position-specific evolutionary preservation. Bioinformatics 32:2230–2232
18. Smigielski EM, Sirotkin K, Ward M, Sherry ST (2000) dbSNP: a database of single nucleotide polymorphisms. Nucl Acids Res 28:52–355
19. Capriotti E, Altman RB, Bromberg Y (2013) Collective judgment predicts disease-associated single nucleotide variants. BMC Genomics 14(Suppl 3):S2
20. Capriotti E, Calabrese R, Casadio R (2006) Predicting the insurgence of human genetic diseases associated to single point protein mutations with support vector machines and evolutionary information. Bioinformatics 22: 2729–2734
21. Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal N (2010) ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. Nucl Acids Res 38:W529–W533
22. Li B, Krishnan VG, Mort ME, Xin F, Kamati KK, Cooper DN et al (2009) Automated inference of molecular mechanisms of disease from amino acid substitutions. Bioinformatics 25(21):2744–2750
23. Källberg M, Wang H, Wang S, Peng J, Wang Z, Lu H (2012) Template-based protein structure modeling using the RaptorX web server. Nat Protoc 7: 1511–1522
24. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y (2015) The I-TASSER Suite: protein structure and function prediction. Nat. Methods 12(1):7–8
25. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ (2015) The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protocols 10: 845–858
26. Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R, Thornton JM (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J Biomol NMR 8:477–486
27. Parthiban V, Gromiha MM, Schomburg D (2006) CUPSAT: prediction of protein stability upon point mutations. Nucl Acids Res 34:W239–W242
28. Zhang Y, Skolnick J (2005) TM-align: a protein structure alignment algorithm based on the TM-score. Nucl Acids Res 33:2302–2309
29. Venselaar H, Te Beek TA, Kuipers RK, Hekkelman ML, Vried G (2010) Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. BMC Bioinformatics 11(1):548
30. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J (2014) STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucl Acids Res 43(Database issue):D447–D452
31. Denver RJ, Bonett RM, Boorse GC (2011) Evolution of Leptin Structure and Function. Neuroendocrinology 94:21–38
32. Prokop JW, Duff RJ, Ball HC, Copeland DL, Londraville RL (2012) Leptin and leptin receptor: analysis of a structure to function relationship in interaction and evolution from humans to fish. Peptides 38(2):326–336
33. Peelman F, VanBeneden K, Zabeau L, Iserentant H, Ulrichts P, Defeau D (2004) Mapping of the leptin binding sites and design of a leptin antagonist. J Biol Chem 279:41038–41046
34. Yadav S, Gupta S, Selvaraj C, Doharey PK, Verma A, Singh SK (2014) *In silico* and *in vitro* Studies on the Protein-Protein Interactions between *Brugia malayi* immunomodulatory protein calreticulin and human C1q. PLOS ONE 9(9):e106413

Publisher's Note