Int J Pharm Bio Sci 2015 July; 6(3): (B) 589 - 599

Research Article **Bioinformatics**

International Journal of Parma and Bio Sciences **ISSN 0975-6299**

IN SILICO ANALYSIS OF FAT MASS OBESITY ASSOCIATED (FTO) GENE USING COMPUTATIONAL ALGORITHMS

PERMENDRA KUMAR, RAJAN KUMAR SINGH AND MAHALINGAM K*

Division of Biomolecules and Genetics, School of Biosciences and Technology VIT University, Vellore-632014, India

ABSTRACT

Non-synonymous single nucleotide polymorphisms (nsSNPs) are used as biomarkers to disease susceptibility. In this study, nsSNPs in Fat Mass Obesity Associated (FTO) gene were screened for its functional impact on the protein. Firstly, SNPs was retrieved from dbSNP database. The computational algorithms namely, PolyPhen and PANTHER were used to identify the potentially deleterious nsSNPs. The following four SNPs, rs139000284, rs139577103, rs368490949 and rs373076420 are found to be most significant SNPs affecting the protein function. The gene network analysis using STRING algorithm showed that the following genes NPY, IGF2BP2, CDKAL1, TCF7L2, SLC30A8, TMEM18, HHEX, TCF7 and GNPDA2 are functionally associated with FTO gene. The homology models of the FTO proteins having the nsSNPs were predicted using SWISS-MODELLER software. The stereochemical properties of the models were checked by Ramachandran plot using PROCHECK algorithm. Root Mean Squared Deviation (RMSD) was calculated by superimposing with native model. The free energy values for these mutant models were higher as compared to their native structure. The present study suggests that these four nsSNPs can be used as biomarkers to screen obesity susceptibility.

KEYWORDS: **Obesity, FTO gene, SNP analysis, Homology Modeling.**

MAHALINGAM K
Division of Biomolecules and Genetics, School of Biosciences and Technology VIT University, Vellore-632014, India

***Corresponding author**

INTRODUCTION

Obesity and overweight represent the major risk factors for the most common chronic illness, including diabetes, cardiovascular diseases and cancer [1]. Once they were considered a problem of developed countries, but overweight and obesity are now considerably increasing in the developing countries. Over the past few years, several genes have been identified from wide-scale screening studies with common variants associated with differences in obesity. More than 50 genes are playing a significant role in the development of obesity. Fat mass and obesity associated (FTO) gene encodes Fe (II) and 2-OG dependent dioxygenase, also known as fat mass and obesity associated (FTO) protein. Recent emerging evidence points to a role of FTO in sensing of nutrients, regulation of translation and growth. A deficiency in FTO leads to postnatal growth retardation and also pointing to some fundamental developmental role. The FTO gene is located on chromosome 16, containing 9 exons having the length of more than 400 Kb nucleotide sequence. It is widely expressed in fetal and adult tissue in human, mice with highest expression in the brain [2,3]. Genome-wide association studies (GWAS) have suggested that FTO gene (fat mass and obesity-associated gene) is the most significant candidate gene contributing to obesity. Recently, it was reported that socio-demographic characteristics and unhealthy eating habits were the major risk factors for the development of childhood overweight and obesity. Sedentary lifestyle along with fast foods, chocolates and fries consumption may lead to top causes for the type 2 diabetes and cardiovascular diseases [4]. Few cases are caused primarily by genes, endocrine disorders, medications or psychiatric illness. It is one of the leading preventable causes of death worldwide. BMI of over 30 is associated with the double mortality rate among women over 16 years period. In United States, It is estimated to cause an excess 111909 to 365000 deaths per year [1]. In Europe, it accounts for 1 million (7.7%) death per year. On average it reduces life expectancy by 6-7 years [5]. BMI of 30-35 reduces life expectancy by 2-4 years. Severe obesity (BMI > 40) reduces life expectancy by 10 years [6]. According to National Family Health Survey (NFHS), the percentage of even married women aged 15-49 years who are overweight or obese increase from 11% in NFSH-2 conducted in the year 1998-99 to 15% in NFSH-3 conducted in 2005-06. In South India the percentage of women who are overweight or obese is higher in Kerala (34%) followed by Tamilnadu (24.4%), Andhra Pradesh (22.7%) and Karnataka (17.3%) is associated with diabetes mellitus, hypertension and breast cancer (NFSH 3, 2005-06). The overall prevalence of overweight and obesity among school going adolescents (10 - 16 years) was found to be 27.8% (overweight – 16.4% and obesity – 11.4%) in eastern India, Bhubaneswar [7]. A study of 38759 European for variant of FTO identified an obesity risk allele [3]. Carrier of one copy of the allelic weighed on average 1.2 Kg more than people with no copies; two copies (16% of the subjects) weighed 3 Kg more and had 1.67 fold higher rate of obesity than those having no copies [3]. Dieting, physical exercise, reducing the energy dense food, intake of fibres is the some physical way to overcome the obesity. Anti obesity drugs are also effective to some extent. Some phytochemicals like *Acanthophora spicifera* (Red algae) extract might be useful. Study has also been reported *Acanthophora spicifera* decreases the level of cholesterol, glucose, protein, creatinine and urea in mice model, where mice have shown a significant decrease in the body weight [8]. The present study was carried out mainly to perform a computational analysis of the SNPs in the FTO gene, to identify the possible disease causing SNPs.

MATERIALS AND METHODS

There are numerous well known computational algorithms that are widely available to predict neutral or deleterious SNPs. In this study we

used the two bioinformatics algorithms, PolyPhen-2 and PANTHER.

(i) Extraction of data sets

The SNPs information of the FTO gene was retrieved from NCBI dbSNP (http://www.ncbi.nlm.nih.gov/snp/) and Swiss-Prot database (http://expasy.org/).

(ii) Simulation for functional change in coding nsSNP by PolyPhen

PolyPhen available from Harvard School of Medicine (http://genetics.bwh.harvard.edu/pph) is a tool which predicts possible impact of amino acid substitution on the structure and function of human proteins based on a combination of phylogenetic, structural and sequence annotation information characterizing a substitution and its position in the protein. PolyPhen input option is a protein sequence or accession number combined with sequence position with amino acid variant [9]. PolyPhen searches for three- dimensional protein structure, multiple alignments of homologues sequence amino acid contact information in various protein databases. Then it calculates position specific independent count (PSIC) score of the two varients. The higher a PSIC score difference, the higher the functional impact a particular amino acid substitution is likely to have. The PolyPhen score can be classified as probably damaging (>2.00), possibly damaging (1.50-1.99), potentially damaging (1.25-1.49), or benign (0.00-0.99).

(iii) Simulation for functional change in coding nsSNP by PANTHER

PANTHER is a database which contains a collection of protein families and subfamilies. It is available free on internet (URL http://www.pantherdb.org/). It predicts how frequently a given amino acid occurs at given position in a family of evolutionary related proteins across different species [10]. The main principle behind working of this tool is Hidden Markov Model (HMM) based statistical modeling method and multiple sequence alignments to perform evolutionary analysis of coding nsSNPs. It calculates subPSEC (substitution position-specific evolutionary conservation score) based on an alignment of evolutionary related proteins and estimates that a particular nsSNP causing a functional impact on the protein. subPSEC score vary from 0 (neutral) to -10 (most likely to be deleterious). The protein sequences having subPSEC score ≤ -3 are said to be deleterious. PANTHER database is maintained by Thomas lab at the University of Southern California.

(iv) Homology Modeling, H-Bond energy and RMSD calculation

Structural stability of native and mutated protein models were performed by structure analysis. We used BLAST to identify the 3D (three dimensional) structure of proteins coded by the genes under study. Also we used homology modeling approach for the prediction of 3D structure. An automated homology modeling program, Swissmodel [11] was performed for modeling. The steps followed in this modeling are as follows: template structure search using BLAST (http://www.ncbi.nlm.nih.gov). The FASTA sequence of FTO was submitted to NCBI BLAST. Following BLAST query, structure of PDB ID 4QKN.1.A (Organism-*Homo sapiens*) was selected as template for FTO protein. A sequence identity of 100% was shown by the template with FTO protein. The target sequence was submitted to Swissmodel automated homology server. The validation for the predicted structure of FTO obtained from Swissmodel was performed by using PROCHECK [12] and the energy minimization performed by Verify3D [13] and NOMAD-Ref server [14]. The overall stereochemistry quality of protein was assessed by Ramchandran plot analysis. NOMAD-Ref server server was used to perform SWISSPDB viewer and energy minimization of mutated protein. Divergence in mutant structure with native structure shows because of mutations and deviation between two structures is elevated by their RMSD. The structures were visualized using Swiss PDB viewer. The H-Bond energy was calculated through VADAR server [15].

(v) STRING (SEARCH TOOL FOR THE RETRIEVAL OF INTERACTING GENES/PROTEINS)

STRING is online tool for the retrieval of interacting Genes/Proteins and available freely worldwide (http://string-db.org/). It is a database of known and predicted protein interactions. The interactions include direct (physical) and indirect (functional) associations. It covers currently above 5.2 millions proteins from 1133 organisms [16].

RESULTS

1. SNP distribution

FTO gene contained total 9460 SNPs of which 61 were nonsynonmous (Source: dbSNP database). We selected all 61 nonsynonymous coding SNPs for our investigation using web based tool i.e., PolyPhen and PANTHER.

2. Prediction of nsSNPs by PolyPhen

The nsSNPs were analyzed for predicting the possible impact of amino acids on the structure and function of the protein using PolyPhen server. The FASTA sequence obtained from NCBI of gene with nsSNP position and their 2 amino acid variants were submitted independently as the input for examine the protein structure change due to amino acid. The result showed that a total of 10 nsSNPs were possibly damaging with PSIC score ≥ 1.0 (Table 1).

Table 1
PolyPhen results of FTO gene.

S. No.	SNP ID	Substitution of nucleotide	Substitution of amino acid	PolyPhen	
				PSIC Score	Predicted Impact
1.	rs77759235	G/T	W383L	1	probably damaging
2.	rs139000284	A/G	R445H	1	probably damaging
3.	rs139577103	A/G	R96H	1	probably damaging
4.	rs140101381	C/T	R80W	1	probably damaging
5.	rs151263395	C/G	P93R	1	possibly damaging
6.	rs182784714	A/C	L146M	1	probably damaging
7.	rs200452822	A/G	E263K	1	probably damaging
8.	rs368490949	C/T	R337C	1	probably damaging
9.	rs370009039	A/T	K216N	1	probably damaging
10.	rs373076420	C/T	P288L	1	probably damaging

3. Prediction of deleterious nsSNP using PANTHER

All the nsSNPs were analyzed using PANTHER for validating its impact on the protein function upon single point mutation. PANTHER can classify proteins by function and hence adding another layer of complexity to refine SNP prediction. PANTHER tool is able to generate multiple outputs, the one of the most useful being the probability that a particular variant is deleterious. A total of 11 nsSNPs were predicted to be deleterious with subPSEC score less than -3 (Table 2).

Table 2
PANTHER results of FTO gene

S. No.	SNP ids	Amino acid change	PANTHER	
			SubPSEC	Prediction
1.	rs76762929	S260C	-3.17229	Deleterious
2.	rs77759235	W383I	-3.15073	Deleterious
3.	rs138348216	M207V	-3.18688	Deleterious
4.	rs139000284	R445H	-4.17883	Deleterious
5.	rs139577103	R96H	-4.65404	Deleterious
6.	rs140101381	R80W	-4.40455	Deleterious
7.	rs149659678	D348V	-3.58468	Deleterious
8.	rs368490949	R337C	-4.79931	Deleterious
9.	rs372814208	P399L	-3.87382	Deleterious`
10.	rs373076420	P288L	-4.93093	Deleterious
11.	rs376381270	R455S	-3.70251	Deleterious

A total of four nsSNPs (rs139000284, rs139577103, rs368490949 and rs373076420) were found to be damaging in causing a change in the structure and function of the protein by both PolyPhen and PANTHER tool (Table 3). Now further studies are focused on these four nsSNPs.

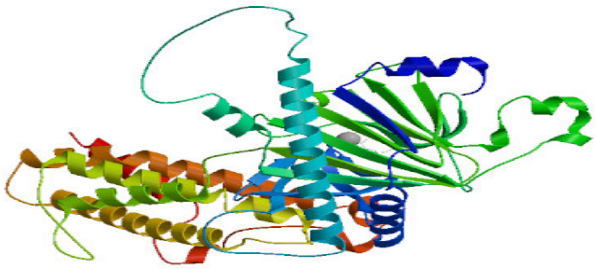
Table 3
List of nsSNPs found to be functionally significant by PolyPhen and PANTHER

S.No.	SNP ID	Nucleotide Substitution	Amino acid Substitution	PolyPhen		PANTHER	
				PSIC Score	Predicted impact	SubPSEC	Prediction
1.	rs139000284	A/G	R445H	1	Probably damaging	-4.1788	Deleterious
2.	rs139577103	A/G	R96H	1	Probably damaging	-4.654	Deleterious
3.	rs368490949	C/T	R337C	1	Probably damaging	-4.7993	Deleterious
4.	rs373076420	C/T	P288L	1	Probably damaging	-4.9309	Deleterious

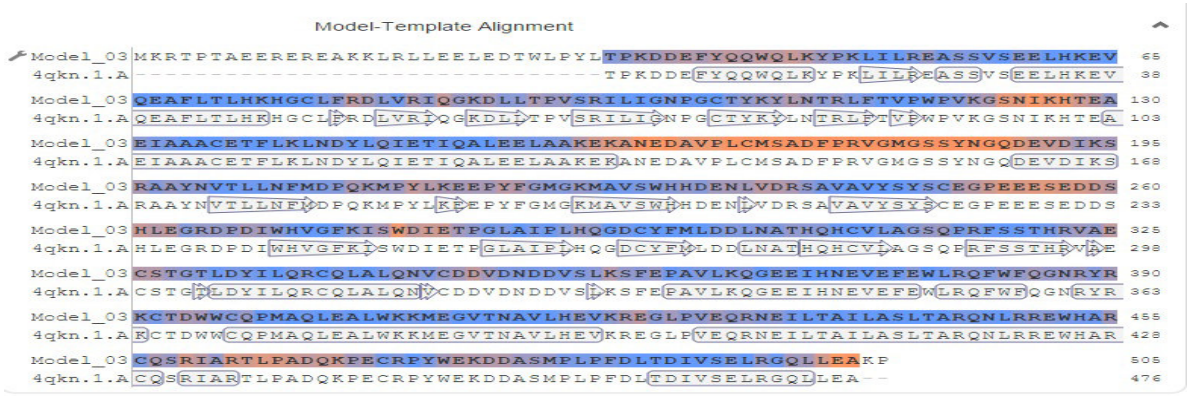
4. Structural modeling of mutant protein

The change in a single amino acid can significantly alter the stability of a protein structure. So, to understand the complete functionality, the knowledge of proteins three dimensional structure is so necessary. Homology modeling method was used to achieve the mapping the damaging nsSNPs into protein structure information. The 3D structure of the predicted models and their selected templates along with its alignment are shown in figure 1. The mutations for the modeled structures (native structure) were performed by SWISS-PDB Viewer and NOMAD-Ref server was used to perform energy minimizations. The mutated structures of rs-ids, rs139000284, rs139577103, rs368490949 and rs373076420 are shown in figure 2. These rs-ids from FTO showed maximum possible result by the servers, PolyPhen and PANTHER. The total energy value of FTO native structure was found to be -21295.650 KJ/mol. The total free energy change, H-Bond energy and RMSD values are summarized in Table 4. for native and mutant models.

Modeled structure of native protein and its alignment



(a)

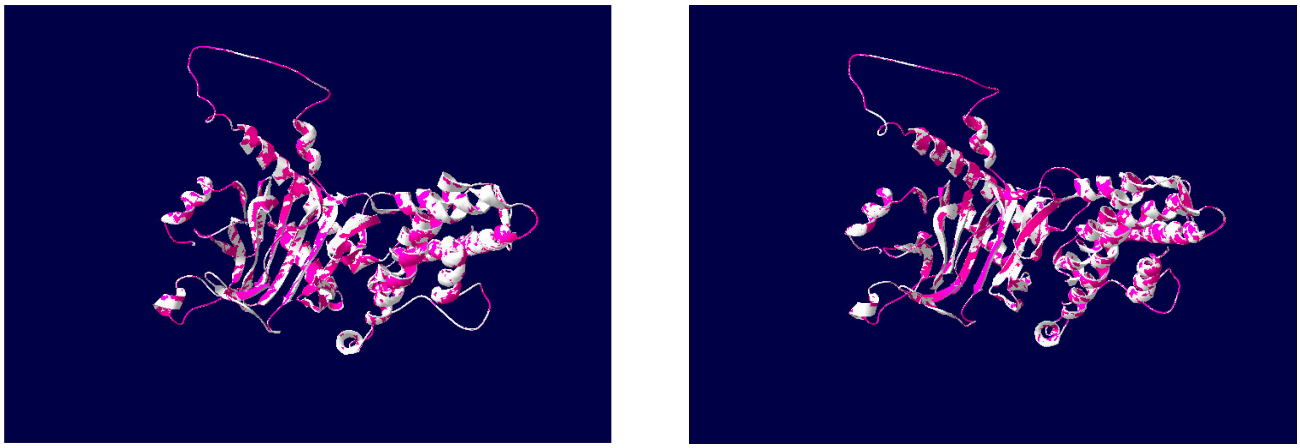


(b)

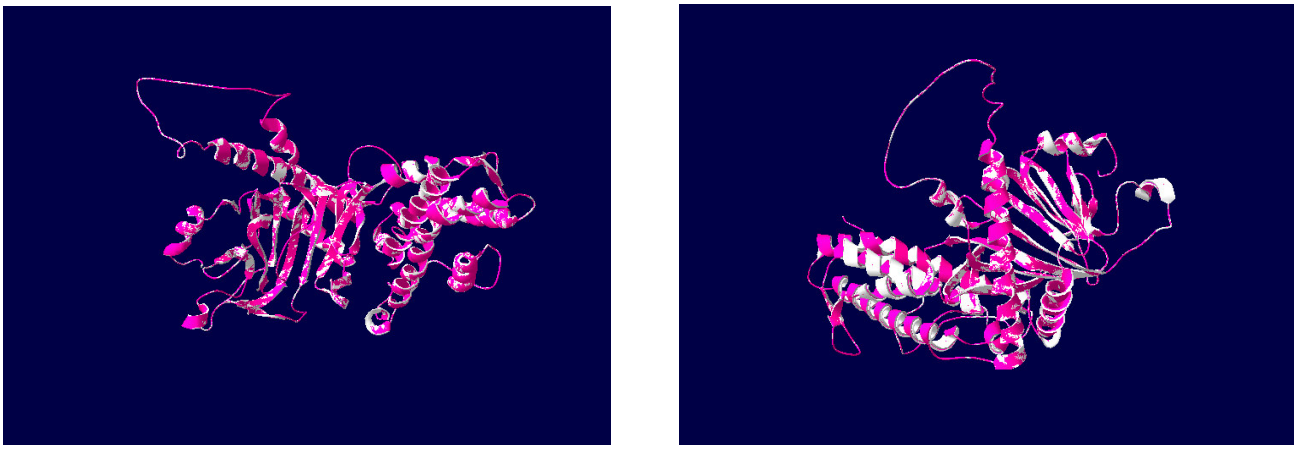
Figure 1

Modeled structure and alignment of FTO native protein (a) Homology modeled 3D structure of native FTO protein (b) PDB ID 4QKN.1.A (Homo sapiens) used as template for FTO protein which shows 100% identity between Model-Template alignments.

Superimposition of native and mutant model



(a) **(b)**



(c) **(d)**

Figure 2
Superimposed structure of native modeled protein with mutant model (a) rs139000284: R445L, (b) rs139577103: R96H, (c) rs368490949: R337C, (d) rs373076420: P228L. White colour showing native structure whereas, pink colour is mutant structure.

Table 4
H bond and total free energy calculation of native and mutated model of FTO protein.

S. No.	SNP ID	H-Bond energy (Native)		H-Bond energy (Mutant)		Total free energy (KJ/mol)		RMSD (Å)
		Donor	Acceptor	Donor	Acceptor	Native	Mutant	
1.	rs139000284	-2.01	-2.00	-1.99	-1.93	-24228.311	-23796.174	0.045
2.	rs139577103	-2.24	-2.57	-2.39	-2.96	-24228.219	-23314.836	0.482
3.	rs368490949	-1.59	-1.85	-1.65	-1.84	-24223.336	-23954.709	0.015
4.	rs373076420	-	-	-	-	-24215.369	-24231.813	0.358

- denotes No prediction

5. PROCHECK Analysis

PROCHECK algorithm was done to check the stereo-chemical properties of the predicted models. It provides the major of structure that how usual the structure is. The G factor value below -0.5 is unusual and below -1 is highly unusual. PROCHECK results are summarized in Table 3. which showed the structures are usual. Therefore, results of PROCHECK depicted that all the structures were found to be usual.

Table 5
PROCHECK results of native and mutated model of FTO protein.

S. No.	rs id	Ramachandran plot statistics				G Factor	
		Most favored region (A, B,L)	Additional allowed region (a,b,l,p)	Generously allowed region [~a,~b,~l,~p]	Disallowed regions (xx)	Average score	Overall average
1	Native structure	88.9%	10.1%	0.7%	0.2%	-0.11	0.07
2	rs139000284	89.2%	9.9%	0.7%	0.2%	-0.11	0.07
3	rs139577103	88.9%	10.1%	0.7%	0.2%	-0.11	0.04
4	rs368490949	88.9%	10.1%	0.7%	0.2%	-0.12	0.06
5	rs373076420	88.9%	10.1%	0.7%	0.2%	-0.11	0.07

6. STRING Analysis

The gene network analysis using STRING algorithm showed that the following genes NPY, IGF2BP2, CDKAL1, TCF7L2, SLC30A8, TMEM18, HHEX, TCF7 and GNPDA2 are functionally associated with FTO gene shown in figure 3.

Gene network analysis

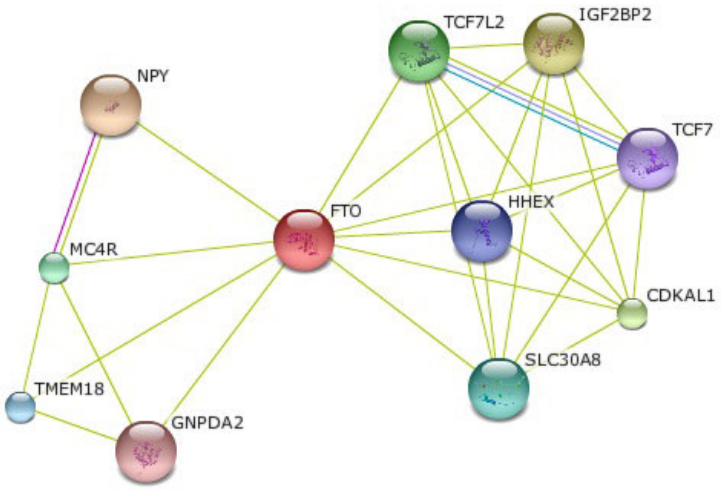


Figure 3
FTO gene having functional interactions with other genes such as NPY, IGF2BP2, CDKAL1, TCF7L2, SLC30A8, TMEM18, HHEX, TCF7 and GNPDA2.

DISCUSSION

SNPs are the most common form of genetic variation among individuals, and accounts for majority of genetic traits as well large number of inherited disease susceptibility. Hence, exploring such SNPs are crucial for detecting its role in disease development. In addition, because of their wide distribution on the species genome, these SNPs turn out to be particularly important and useful genetic markers in the research .The PolyPhen web server designed by Bork and coworkers [9] uses both sequence and structural information to forecast deleterious and non-deleterious mutations. BLAST is used to create a multiple sequence alignment from homologous proteins (30–94% sequence identity with the target), and PSIC [17] is used to construct a position-specific scoring matrix (PSSM) from this alignment. PolyPhen uses BLAST to find proteins of identified structure homologous to the target sequence, but limits its results in the cases for which the amino acid in the structure is the same as the wild-type amino acid under study. PolyPhen uses the sequence alignment and the identified structure to locate residue accessibility and proximity to ligands and interacts with other subunits in the structure. A total of 14 nsSNPs were predicted to damaging by PolyPhen in this study. Protein structural analysis was carried out based on the screened results obtained from PolyPhen and PANTHER. Protein 3D structural information is a key feature for predicting the effects of deleterious nsSNPs and protein structure analysis gives information about the environment of the mutation. Proteins with mutations do not always have 3D structures that are analyzed and deposited in PDB. Therefore, it is necessary to construct 3D models using molecular modeling protocols [11]. This is a simple way of detecting what kind of adverse effects that a mutation can have on a protein. Single amino acid

substitutions were studied using Swiss-PDB viewer with RMSD values calculated from mutant and native modeled structures [18]. Computing the energy gives the information about the protein structure stability. The total energy values of native and mutated modeled structures were compared for the genes under study. Mutant structures of homology modeled FTO at position R445H, R96H, R337C and P288L showed an increase in total energy level (less favorable change) and increase in RMSD value deviation in comparison with native structure. To improve the strength of our analysis, we analyzed the data by the combinations of all the tools used. Significant concordance was observed between the functional consequences of nsSNPs. By comparing the scores of all the methods used in this analysis, four nsSNPs with IDs rs139000284, rs139577103, rs368490949 and rs373076420 of FTO were predicted to be functionally significant. Single nucleotide polymorphism (SNP) is one the most commonly known genetic variation which is used to map complex genetic traits. The database of SNP, dbSNP contains a lot of SNPs distributed across the genes [19]. The nsSNPs are affecting gene expression by changing DNA and transcription factor binding [20] and ultimately inactivate the active sites of enzymes or alter the splice site, which results into the production of defective gene products [21]. Epidemiologic association studies put a great amount of effort in finding SNPs in the genes that may have associated with disease susceptibility. Numerous molecular epidemiologic studies focus on SNPs found in coding regions in hope of finding significant association between SNPs and diserease risk, but often find little or no association [22]. The population of snSNPs is also growing rapidly with the availability of high-throughput of SNP detection techniques. It provides a platform for studying the association between genotype and phenotype of human disease. The selection of functionally polymorphic nsSNP for an association study can be boosted by investigating the possible effect an amino acid variant may have on the function of the encoded protein with the help of different SNP detection tools like I- Mutant, Sort Intolerant from Tolerant (SIFT) and Polymorphism Phenotype (PolyPhen) etc [22]. Discovering the deleterious nsSNPs out of a group of all SNPs will be very valuable for epidemiological population based studies. Hence, the current study was undertaken to spot the potential deleterious nsSNPs in FTO gene known for their involvement in obesity susceptibility.

CONCLUSION

From the results obtained by computational algorithms it is conclude that nsSNP with rs ids rs139000284, rs139577103, rs368490949 and rs373076420 were found to be probably damaging by PolyPhen tool with PSIC score was 1 for each and deleterious by PANTHER tool with SubPSEC score was > - 4. The free energy of each mutant model was found to be more than that of native one. RMSD value for the following two rs ids, rs139577103 and rs373076420 are found to be significantly crucial. The H-bond energy of these mutants is also found to be slightly different from native one. By using STRING tool the interaction of FTO with other gene has been shown. The results from these various computational tools suggest that the above mentioned four nsSNPs may be suspected for the development of obesity . The in-vitro and population based studies need to be done to validate this results.

ACKNOWLEDGEMENT

Authors are grateful to VIT University, Vellore, Tamilnadu, India, for providing facility for this work. Permendra Kumar is also thankful to UGC for providing RGNF-JRF for the presented work.

CONFLICT OF INTEREST

Authors having no conflict of interest for publishing of this work.

REFERENCES

1. Haslam DW, James WP, Obesity. Lancet, 366 (9492): 1197–209, (2005).
2. Vaisse C, Clement K, Guy-Grand B, Froguel PA, frameshift mutation in human MC4R is associated with a dominant form of obesity. Nature Genetics, 20 (2): 113–114, (1998).
3. Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, Lindgren CM, Perry JRB, Elliott KS, Lango H et al., A common variant in the FTO gene is associated with body mass index and predispose to childhood and adult obesity. Science 316 (5826): 889-894, (2007).
4. Manasi Behera, Dipti Mohapatra, Nibedita Priyadarsini, Tapaswini Mishra and Priyambada Panda, Prevalence and risk factors of childhood overweight and obesity. International Journal of Pharma and Bio Sciences, 6 (1): (B) 91 – 95, (2015).
5. Peeters A, Barendregt JJ, Willekens F, Mackenbach JP, Al Mamun A and Bonneux L, Obesity in adulthood and its consequences for life expectancy: A life-table analysis. Annals of Internal Medicine, 138 (1): 24–32, (2003).
6. Whitlock G, Lewington S, Sherliker P, Clarke R, Emberson J, Halsey J, Qizilbash N, Collins R, Peto R and 157 prospective studies collaborators, Body-mass index and cause-specific mortality in 900000 adults: collaborative analyses of 57 prospective studies. Lancet, 373 (9669): 1083-96, (2009).
7. Lipilekha Patnaik, Sumitra Pattanaik, E Venkata Rao and Trilochan Sahu, Overweight/ Obesity among school going adolescents and its associated Factors - A cross sectional study in an urban locality of eastern india. International Journal of Pharma and Bio Sciences, 6 (2): (B) 1252 – 1258, (2015).
8. K Mohana Priya and Samanta S Khora, Ameliorative effect of red algae A*canthophora spicifera* against puffer fish *Arothron hispidus* induced biochemical and oxidative stress in mice. International Journal of Pharma and Biosciences, 4(3): 1043 – 1052, (2013).
9. Ramensky V, Pork P and Sunyaev S, Human non-synonymous SNPs: server and survey. Nucleic Acids Research, 30 (17): 3894–3900, (2002).
10. Mi H, Guo N, Kejariwal A and Thomas PD, PANTHER version 6: protein sequence and function evolution data with expanded representation of biological pathways. Nucleic Acids Research, 35: D 247–252, (2007).
11. Arnold K, Bordoli L, Kopp J and Schwede T, The SWISS-MODEL workspace: a web based environment for protein structure homology modeling. Bioinformatics, 22 (2): 195–201, (2006).
12. Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R and Thornton JM, AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. Journal of Biomolecular NMR, 8 (4): 477–486, (1996).
13. Bowie JU, Luthy R and Eisenberg D, A method to identify protein sequences that fold into a known three-dimensional structure. Science, 253 (5016): 164–170, (1991).
14. Lindahl E, Azuara C, Koehl P and Delarue M, NOMAD-Ref: visualization, deformation and refinement of macromolecular structures based on all-atom normal mode analysis. Nucleic Acids Research, 34: W52–W56, (2006).
15. Willard L, Ranjan A, Zhang H, Monzavi H, Boyko RF, Sykes BD and Wishart DS, VADAR: a web server for quantitative evaluation of protein structure quality. Nucleic Acids Research, 31 (13): 3316–3319, (2003).
16. Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, Lin J, Minguez P, Bork P, Von Mering C and Jensen LJ, STRING v9.1: protein-protein interaction networks with increased coverage integration. Nucleic Acids Research, 41: D808 - D815, (2013).

17. Sunyaev S, Hanke J, Aydin A, Wirkner U, Zastrow I, Reich J and Bork P, Prediction of non-synonymous single nucleotide polymorphisms in human disease-associated genes. Journal of Molecular Medicine, 77 (11): 754–760, (1999).
18. Guex N and Peitsch MC**,** SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. Electrophoresis*,* 18 (15)**:** 2714-2723, (1997).
19. Smigielski EM, Sirotkin K, Ward M and Sherry TS, dbSNP: a database of single nucleotide polymorphisms. Nucleic Acids Research, 28 (1): 352-355, (2000).
20. Barroso I, et. al., Dominant negative mutations in human PPAR gamma associated with severe insulin resistance, diabetes mellitus and hypertension. Nature, 402 (6764): 880–883, (1999).
21. Jaruzelska J, Abadie V, d Aubenton-Carafa Y, Brody E, Munnich A and Marie J, In vitro splicing deficiency induced by a C to T mutation at position-3 in the intron 10 acceptor site of the phenylalanine hydroxylase gene in a patient with phenylketonuria. Journal of Biological Chemistry, 270 (35): 20370–20375, (1995).
22. Johnson MM, Houck J and Chen C, Screening for deleterious nonsynonymous single-nucleotide polymorphisms in genes involved in steroid hormone metabolism and response. Cancer Epidemiology, Biomarkers and Prevention, 14 (5): 1326–1329, (2005).