Cellular and Molecular Biology
E-ISSN : 1165-158X / P-ISSN : 0145-5680
www.cellmolbiol.org

Computational analysis of missense variants of human MC4R and childhood obesity

Imane Douiyeh[1,2*], Jihane Khamlich[1,2], Asmae Saih[2], Asmae Baggar[1], Anass Kettani[2,3#], Amal Safi[1#]

[1] Laboratory Biochemistry Environment and Agri-food, Department of Biology, Faculty of Science and Technics Mohammedia, Hassan II University Casablanca, Morocco
[2] Laboratory of Biology and Health, URAC 34, Faculty of Sciences Ben M'Sik Hassan II University of Casablanca, Morocco
[3] Mohammed VI Center for Research & Innovation, Rabat, Morocco & Mohammed VI University of Health Sciences, Casablanca, Morocco

[#]These authors contributed equally to this work

ARTICLE INFO

Original paper

Article history:
Received: April 13, 2023
Accepted: July 05, 2023
Published: October 31, 2023

Keywords:

MC4R, childhood obesity, SNP, prediction, in-silico analysis, pathogenicity

ABSTRACT

Industrialized and developing nations face severe public health problems related to childhood obesity. Previous studies revealed that the melanocortin-4 receptor gene (MC4R) is the most prevalent monogenic cause of severe early obesity. Due to its influence on food intake and energy expenditure via neuronal melanocortin-4 receptor pathways, MC4R is recognized as a regulator of energy homeostasis. This study used a variety of computational systems to analyze 273 missense variations of MC4R in silico. Several tools, including PolyPhen, PROVEAN, SIFT, SNAP2, MutPred2, PROVEAN, SNP&GO and Mu-Pro, I-Mutant, PhD-SNP, SAAFEC-SEQ I-Mutant, and ConSurf, were used to make predictions of 13 extremely confident nsSNPs that are harmful and disease-causing (E308k, P299L, D298H, C271F, C271R, P260L, T246N, G243R, C196Y, W174C, Y157S, D126Y, and D90G). The results of our study suggest that these MC4R nsSNPs may disrupt normal protein function, leading to an increased risk of childhood obesity. These results highlight the potential use of these nsSNPs as biomarkers to predict susceptibility to obesity and as targets for personalized interventions.

Doi: http://dx.doi.org/10.14715/cmb/2023.69.10.5

Introduction

Childhood obesity is recognized as an important public health problem in industrialized countries. Obese people are subject to many difficulties in developing countries (increased mortality, metabolic, and cardiovascular diseases) (1). The leading cause of the current global obesity epidemic is the evolution of our lifestyle, which is considered an obesogenic environment (easy access to high-calorie foods combined with reduced physical activity). 70% of polygenic obesity cases are caused by weight heritability, and 5% of monogenic obesity cases are monogenic due to a single genetic mutation in the leptin-melanotic pathway, which is satiety (2). So far, nine genes (MC4R, NTKR2, SH2B1, BDNF, LEP, POMC, LEPR, PCSK1, SIM1) have been identified (3).

It is essential to understand the molecular processes underlying the accumulation and expenditure of energy and the regulation of food intake without forgetting the defects of these regulations that lead to obesity.

The hypothalamus regulates digestion and metabolism in the arcuate nucleus (ARC). Two main pathways control them. The first is the anorexigenic signaling pathway, which is mediated by leptin which reduces food intake while increasing energy expenditure via the neurons POMC (Proopiomelanocortin) and the CART transcript. The second is the orexigenic signaling pathway which is, this time, mediated by ghrelin. The latter increases food intake while decreasing energy expenditure via neuropeptide Y (NPY) and AgRP. This study focused on the orexigenic signaling pathway, specifically the MC4R protein, which is activated by melanocortin-stimulating hormone alpha (-MSH) and produced by POMC neurons (Figure 1) (4,5).

The melanocortin-4 receptor *(MC4R)* gene is the most

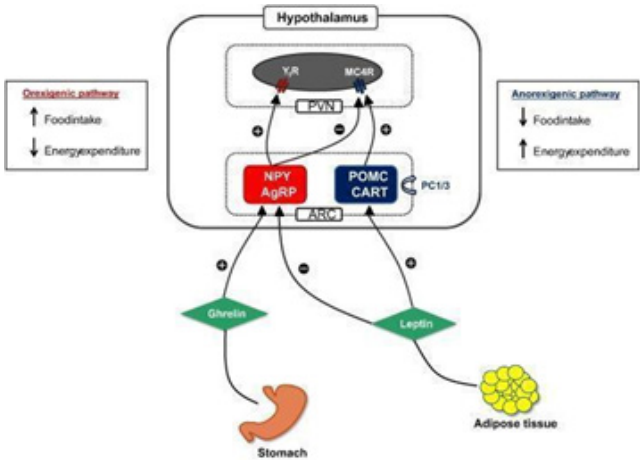


Figure 1. The hypothalamic leptin-melanocortin pathway (4): Two major proteins regulate the control of appetite: leptin and ghrelin, which refer to the hypothalamic leptin-melanocortin pathway. The diagram represents 2 pathways, the anorectic pathway, which is activated by leptin and prevents food intake, and the orexigenic pathway, which is activated by ghrelin and promotes food intake.

* Corresponding author. Email: imane.douiyeh-etu@etu.univh2c.ma

common monogenic form of early-onset severe obesity and is considered a regulator of energy homeostasis and impacts food intake and energy consumption via neuronal melanocortin-4 receptor signaling (6,7).

The most common form of genetic variability in humans is single nucleotide polymorphism, which accounts for about 1/1,000th of the average human genome. As more genomes have been sequenced over the past few years, our understanding of genetic and genomic diversity has grown exponentially. As a result, the relative costs of sequencing and genotyping technologies have decreased. Parallel to this development, the quantity, size, and scope of bioinformatics databases and software related to the gathering and processing of genetic data have expanded, as have the discoveries of genes causing monogenic and complicated disorders (8).

The mechanisms by which a nucleotide variant may induce phenotypic change are primarily unknown. *In-silico* analysis using performant software makes it possible to predict the phenotypic effect of non-synonymous coding SNPs on the physico-chemical properties of the proteins involved. Providing this data is essential for understanding disease biology and genotype-phenotype connections. This bioinformatic study may deepen our comprehension of the genetics underlying obesity by identifying and characterizing deleterious mutations in the *MC4R* gene, which may improve our knowledge of the regulation of energy balance and the identification of patients who may benefit from treatment with an MC4R agonist.

In this project, we sought to bioinformatically examine, via different tools that study the pathogenicity, stability and flexibility of the protein, the current evidence for genetic associations of childhood obesity with the MC4R gene in order to highlight and identify the nsSNPs involved, which will make it easier in terms of cost and efficiency for researchers to pursue the work experimentally.

Materials and Methods

The schematic illustration of the methodology utilized in this study is given in Figure 2.

Data retrieval

The Ensembl Genome Browser (http://www.ensembl.org) was used to extract SNP data (rsID, chromosomal location, position and residue change) from the MC4R gene, of which our exact target was nsSNPs. The MC4R variants were studied in depth to determine their effects on the protein.

For this purpose, Uniprot (https://www.uniprot.org/uniprot/K4MTP3.fasta) was used to extract the amino acid (AA) sequence of the gene, to use it to study the structural and functional impact of the selected nsSNPs on the protein.

Prediction of damaging nsSNPs of the *MC4R* gene

We used several tools to extrapolate the effects of non-synonymous SNPs in the *MC4R* gene on protein function, including Single Nucleotide Polymorphisms & Gene Ontology **(SNPs&GO)** (http://snps.biofold.org/snps-and-go/snps-and-go.html), Sorting Intolerant From Tolerant **(SIFT)**(http://blocks.fhcrc.org/sift/SIFT.html)**,** Protein Variation Effect Analyzer **(PROVEAN)** (http://provean.jcvi.org/index.php), Polymorphism Phenotyping **(PolyPhen)** (http://genetics.bwh.harvard.edu/pph/), and Predictor of human Deleterious Single Nucleotide Polymorphisms**(PhD-SNP)** (http://snps.biofold.org/phd-snp/phd-snp.html).

SIFT is an online server that takes as input a target protein sequence and uses various alignment information to predict the nature of the amino acid substitutions, whether benign or deleterious, for each position of this sequence (score ≤ the algorithm predicts 0.05 to be damaging to varying degrees, while a score > 0.05 is designated as tolerant (9).

PolyPhen is a web server that allows probable categorization by generating predictions that characterize substitutions at the protein level based on sequence, structural, and phylogenetic annotation information. This program works by giving a pathogenicity classification to substitutions (10).

SNAP2 is a machine-learning server that predicts the functional effect of an amino acid substitution present in a given protein. A distinction is made between variants with a deleterious effect and those with a benign effect. As an output, the analysis generates prediction scores that indicate whether or not there is a correlation between the severity of the effect and the prediction score (11).

PROVEAN is an in-silico analyzer that provides an approach for predicting the functional effects of single and multiple amino acid substitution. Based on alignment, it generates scores that allow the classification of mutations as deleterious or not (12).

SNPs&GO is a support vector machine-based database that uses prediction scores to distinguish deleterious effects from single-position polymorphisms (13).

PhD-SNP is a predictor that provides an analysis of the changes generated by substitutions to determine whether these SNPs are associated with pathogenicity or not (14).

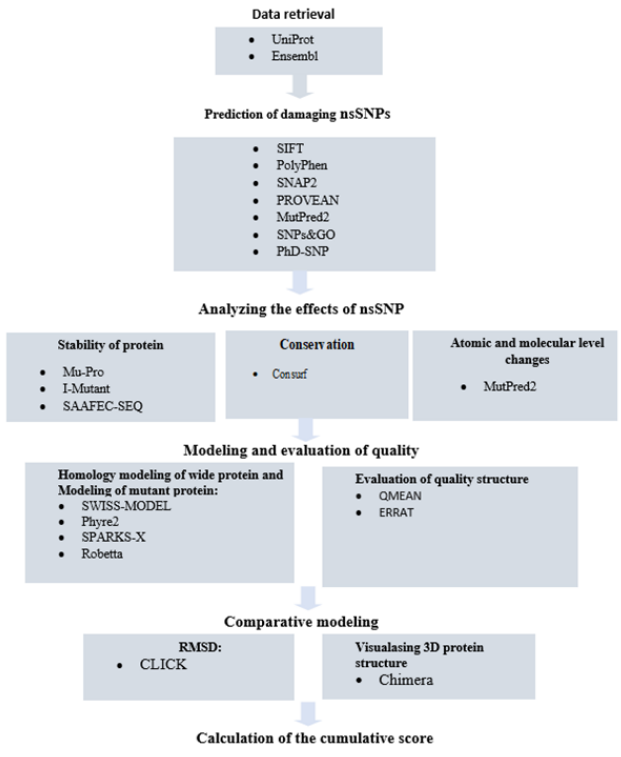


Figure 2. Schematic representation of the overall schema followed in the study.

Analyzing the effect of nsSNPs

Prediction of the effects of nsSNPs on protein stability

SNPs typically has an effect on protein stability by either decreasing or increasing protein strength. However, it is crucial to accurately forecast the changes in protein stability brought on by single amino acid alterations. We used three tools to predict these effects and enhance our confidence considering the changes they brought about.

MUpro is a web server that determines how amino acid mutations affect the various protein states and strengths (http://mupro.proteomics.ics.uci.edu) (http://mupro.proteomics.ics.uci.edu) (15).

I-Mutant: It can forecast changes in protein stability when there are variations at a single site in the structure or sequence of the protein. The web server result is created in the following way: DDG = 0: neutral, DDG > 0: increased stability, or DDG 0: decreased stability (http://gpcr2.biocomp.unibo.it/cgi/predictors/I-Mutant3.0/I-Mutant3.0.cgi) (16).

SAAFEC-SEQ is an online tool for determining how missense mutations affect the folding free energy of proteins (http://compbio.clemson.edu/SAAFECSEQ/index.php).

Conservation of amino acid positions

The stored-product acids amines are the ones most likely to be functionally significant. Future studies of molecular epidemiology may find it useful to use the evolutionary molecular approach as a powerful tool for hierarchizing SNPs. Polymorphisms that alter these amino acids may therefore have a higher propensity to be linked to disease vulnerability.

Consurf is a tool based on the phylogenetic analysis between homologous sequences to determine the evolutionary conservation of amino acid locations (http://ConSurf.tau.ac.il/). The conservation analysis is displayed for each residue of the target protein on a scale from 1 to 9. According to the scale, regions with scores of 1-3 are described as changeable, 4-6 as average, and 7-9 as highly conserved (17).

Atomic and molecular level changes

We used the MutPred server (http://mutpred.mutdb.org/) to predict the atomic and molecular level changes of the disease-related substitution.

MutPred2 is based on structure and sequence changes at functional sites between wild-type and mutant sequences. Scores are assigned for each gain or loss of structure and function, which can provide insight into the precise molecular mechanism of the disease state. If the overall MutPred score is greater than 0.5 and the property P-score is less than 0.05, the server considers a prediction of the underlying disease mechanism as a workable hypothesis (http://mutpred.mutdb.org) (13).

Modeling and evaluation of quality

The functional study of a protein requires structural data obtained by NMR or RX. Otherwise, it will be necessary to use homology modeling. Homology modeling starts from the hypothesis that proteins with homologous sequences have closely related structures (18). Indeed, it is a comparative modeling technique that requires a priori knowledge of the3D structure of a protein called a template.

Homology modeling

We used several structural bioinformatics web servers dedicated to homology modeling to create 3D structures of the native protein and the MC4R mutants to estimate the deleterious effects of the chosen nsSNPs on the protein structure, including Robetta(https://robetta.bakerlab.org/ SWISS-MODEL(https://swissmodel.expasy.org/), RAPTORX (http://raptorx.uchicago.edu/), Phyre2 (http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index), SPARKS-X (https://sparks-lab.org/server/sparks-x/).

Evaluation of quality structure

In order to compare and choose the model with the best quality to get better results, we verified the quality of the models using the Qualitative Model Energy Analysis (QMEAN) server and the ERRAT server.

QMEAN estimates the "degree of nativeness" of the model's structural characteristics worldwide. It compares the QMEAN score of the model structure with the score of a comparable-sized experimental structure. If there is a solid accord between the experimental structure and the model structure, the Z-scores are around zero. On the other hand, a low-quality model has scores lower or equal to -4.0(https://swissmodel.expasy.org/qmean/help) (19).

ERRAT, based on a distinctive atomic interaction, distinguishes between portions of proteins that have been correctly and erroneously defined(http://services.mbi.ucla.edu/ERRAT/)(20).

Comparative modeling

Visualization of 3D protein structures

Chimera is a molecular graphics software that has been developed to enhance interactive visualization, analytic systems, and related data (21). This program was utilized to generate high-quality images of both native and mutant residues for the purpose of studying mutations (http://www.cgl.ucsf.edu/chimera/).

Comparison of native and mutant protein structures

CLICK is a tool designed to compute a comparison of the structural differences between wild-type and mutant structures (http://cospi.iiserpune.ac.in/click/#). The tool computed RMSD (the distance between the two structures after superimposition) and a Z-score (the significance of the alignment) for each alignment (22).

Protein flexibility

Predyflexy is designed to predict both flexibility of proteins and their local structure. The results are provided in two different formats: a graph displays the sequence based on the predicted flexibility (ranging from 0 to 2 for highly flexible residue) and the confidence index (ranging from 1 to 19, with increasing values reflecting higher prediction accuracy), as well as a table containing detailed information on each amino acid in the target protein's sequence (https://www.dsimb.inserm.fr/dsimb_tools/predyflexy/index.html)(23).

DynOmics ENM

DynOmics ENM is a robust tool used for determining dynamic cross-correlation maps (DCCM). The server is based on two elastic networks models (ENMs): the Gaussian Network Model (GNM) and the Anisotropic Network Model (ANM). It combines three distinct components: an

assessment of the collective motions of the biomolecules, an evaluation of the critical sites involved in the chemical process, and the resolution exchanges between the full atomic and CG representations (http://enm.pitt.edu/) (24).

Calculation of the cumulative score

To predict the high-risk pathogenic nsSNPs of the human *MC4R gene*, we calculated the cumulative score to gather all the software tools that were used for each nsSNP (SIFT, polyphen, SNAP2, PROVEAN, Mutpred2, SNPs&GO, PHD-SNP, ConSurf, I-Mutant, MUpro, SAFEK-SEQ, and CLICK), and then a value was calculated using the Sum function of Excel. We will be able to predict the most harmful nsSNPs by using a score that takes into account the results of all the servers we have utilized, which in turn enhances the reliability of our predictions.

Results

Data retrieval

Within the exonic region of the human MC4R gene, there are a total of 731 variants. Among these variants, 518 are classified as single nucleotide polymorphisms (SNPs), while 273 are reported to be missense variants.

Prediction of damaging SNPs

SIFT, SNAP2, PROVEAN, PolyPhen, MutPred2, SNP&GO and PhD-SNPservers predicted that 16 missense variants in the exonic region of the *MC4R* gene were deleterious (Table 1).

nsSNPs effects evaluation

Atomic and molecular level changes of MC4R by MutPred server

The estimated structural and functional changes consist of gains in sulfatation and N-linked glycosylation, losses in sulfatation and N-linked glycosylation, and altered metal binding. Their P values are provided in Table S1.

Any analysis of the disease's fundamental etiology presented in the Table S1 has a general score above 0.5 and an individual property score of $P < 0.05$. Thus, we considered all predictions to be confident. Except for the C271R and G243 mutations, whose prediction was viewed as having a high degree of confidence (general score corresponds to a specificity of 0.95 and the gain/loss property score $P < 0.01$).

Protein stability effect prediction (T=25, pH=7)

Three servers were used to predict the potential structural impact of candidate nsSNPs. According to I-Mutant, two nsSNPs, H76Y and N62S, out of 16 variants were found to increase the stability of the protein. SAAFEC-SEQ predicted only one nsSNP T150I, which has a stabilizing effect on the protein. While MuPro predicted that all nsSNPs decrease the stability of the protein. Results revealed that 13 nsSNPsdestabilized and decreased the stability of the MC4R protein. Table 3 presents an illustration of the outcomes.

The results obtained by ConSurf show the structural and functional conservation levels of all MC4R residues. The sequence conservation study was performed on the 13 previously selected nsSNPs. E308K, G243R, D126Y, and D90G were suggested by the data to be functional residues, which makes them highly exposed and conserved. The structural residues P299L, D298H, C271F, C271R, P260L, T246N, W174C, and Y157Sare projected to be deeply buried and well-preserved in Table 4 and Fig 3.

Modeling and quality checking

Homology modeling

The Table 5 summarizes the results of the prediction of a 3D structure with 4 servers and the results of evaluating the quality of the predicted structures. For the rest of the work, we chose the model generated by the SWISS-MODEL server, which represents the highest identity percentage (100%) and good quality (QMEAN: 0.79, ERRAT = 99.2509) (Table 5).

Table 1. Predicted functional nsSNPs in *MC4R* by seven servers.

Variant ID	SIFT		PolyPhen		SNAP2		PROVEAN		MutPred2		SNP&GO		PhD-SNP	
	Pred	Sc	Pred	Sc	Pred	Sc	Sc	Pred	Pred	Sc	Pred	RI	Pred	RI
rs375095163	Del	0	Prb.damag	0.999	effect	37	-3.281	Del	Path	0.882	Dis	0	Dis	7
rs52804924	Del	0	Prb.damag	1	effect	64	-9.027	Del	Path	0.839	Dis	2	Dis	7
rs760199460	Del	0	Prb.damag	0.999	effect	69	-6.188	Del	Path	0.874	Dis	1	Dis	5
rs121913562	Del	0	Prb.damag	1	effect	80	-10.660	Del	Path	0.927	Dis	1	Dis	9
rs1057517991	Del	0	Prb.damag	1	effect	78	-11.629	Del	Path	0.952	Dis	1	Dis	9
rs1435358988	Del	0	Prb.damag	1	effect	74	-9.700	Del	Path	0.843	Dis	2	Dis	8
rs1333658154	Del	0	Prb.damag	0.999	effect	61	-4.899	Del	Path	0.885	Dis	0	Dis	7
rs868309222	Del	0	Prb.damag	0.999	effect	87	-7.813	Del	Path	0.956	Dis	1	Dis	7
rs1191554117	Del	0	Prb.damag	1	effect	90	-9.587	Del	Path	0.914	Dis	2	Dis	5
rs1159323398	Del	0	Prb.damag	0.998	effect	66	-12.394	Del	Path	0.946	Dis	3	Dis	8
rs768916374	Del	0	Prb.damag	1	effect	88	-8.844	Del	Path	0.915	Dis	2	Dis	8
rs766665118	Del	0	Prb.damag	1	effect	45	-5.376	Del	Path	0.834	Dis	3	Dis	6
rs768806551	Del	0	Prb.damag	1	effect	83	-8.860	Del	Path	0.925	Dis	1	Dis	5
rs1598932350	Del	0	Prb.damag	1	effect	91	-6.681	Del	Path	0.918	Dis	1	Dis	7
rs1215552316	Del	0	Prb.damag	1	effect	63	-5.576	Del	Path	0.877	Dis	2	Dis	6
rs121913566	Del	0	Prb.damag	1	effect	95	-4.635	Del	Path	0.758	Dis	2	Dis	5

Legend: Sc: Score, Pred: Prediction, Del: Deleterious, Effect: Pathogenic effect, Prb.damag: probably damaging, Path: Pathogenic, Dis: Disease.

Table 3. Prediction of MC4R protein stability.

Variant ID	Location ch8	mutation	I-mutant		Mu-Pro		SAAFEC-SEQ	
			Stability	DDG	Stability	DDG	Stability	DDG
rs375095163	60371428	E308K	Decrease	-0.55	Decrease	-0.78	Destabilizing	-0.27
rs52804924	60371454	P299L	Decrease	-0.95	Decrease	-0.40	Destabilizing	-0.43
rs760199460	60371458	D298H	Decrease	-0.73	Decrease	-1.55	Destabilizing	-0.20
rs121913562	60371538	C271F	Decrease	-0.52	Decrease	-0.63	Destabilizing	-1.08
rs1057517991	60371539	C271R	Decrease	-0.90	Decrease	-1.04	Destabilizing	-1.21
rs1435358988	60371571	P260L	Decrease	-1.03	Decrease	-0.06	Destabilizing	-0.23
rs1333658154	60371613	T246N	Decrease	-0.49	Decrease	-1.19	Destabilizing	-0.37
rs868309222	60371623	G243R	Decrease	-1.76	Decrease	-0.61	Destabilizing	-0.30
rs1191554117	60371763	C196Y	Decrease	-0.69	Decrease	-1.19	Destabilizing	-1.03
rs1159323398	60371828	W174C	Decrease	-1.05	Decrease	-0.46	Destabilizing	-0.70
rs768916374	60371880	Y157S	Decrease	-2.53	Decrease	-1.44	Destabilizing	-1.84
rs766665118	60371901	T150I	Decrease	-0.41	Decrease	-0.94	stabilizing	0.09
rs768806551	60371974	D126Y	Decrease	-1.91	Decrease	-0.77	Destabilizing	-0.25
rs1598932350	60372081	D90G	Decrease	-3.27	Decrease	-1.84	Destabilizing	-0.44
rs1215552316	60372124	H76Y	Increase	1.17	Decrease	-0.38	Destabilizing	-0.25
rs121913566	60372165	N62S	Increase	0.87	Decrease	-1.28	Destabilizing	-0.64

Sequence conservation analysis.

Table 4. Sequence conservation analysis.

Variant ID	Location ch8	SNP	Normalized score	score	Conservation Degree
rs375095163	60371428	E308K	-0.808	9	Highly conserved and exposed
rs52804924	60371454	P299L	-0.804	9	Highly conserved and buried
rs760199460	60371458	D298H	-0.809	9	Highly conserved and buried
rs121913562	60371538	C271F	-0.798	9	Highly conserved and buried
rs1057517991	60371539	C271R	-0.798	9	Highly conserved and buried
rs1435358988	60371571	P260L	-0.804	9	Highly conserved and buried
rs1333658154	60371613	T246N	-0.812	9	Highly conserved and buried
rs868309222	60371623	G243R	-0.472	8	Highly conserved and exposed
rs1191554117	60371763	C196Y	-0.539	8	Highly conserved and buried
rs1159323398	60371828	W174C	-0.781	9	Highly conserved and buried
rs768916374	60371880	Y157S	-0.800	9	Highly conserved and buried
rs768806551	60371974	D126Y	-0.774	9	Highly conserved and exposed
rs1598932350	60372081	D90G	-0.809	9	Highly conserved and exposed

Visualization of 3D protein structures

We first obtained the 3D structures of the wild-type protein and 13 mutants using SWISS-MODEL in order to ascertain if the 13 high-risk nsSNPs affect the MC4R protein's wild-type structure. The MC4R wild-type sequence had each nsSNP manually replaced, and the sequences were then sent into the SWISS-MODEL homology modeling program. SWISS-MODEL used (PDB ID 7f53.1. E) as the template for predicting 3D models. Subsequently, we validated and confirmed the models using ERRAT, in addition to output data obtained from the SWISS-MODEL Ramachandran plot, QMEAN score, and GMQE score. This allowed us to evaluate the overall quality of the generated models, which is presented in Table 6. The structures were visualized using Chimera (Figure 4).

Comparative modeling

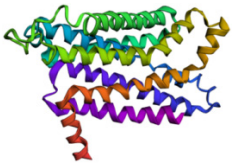
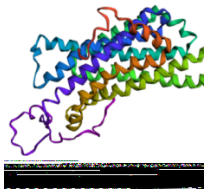
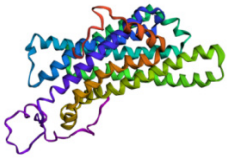
An analysis of the structural differences between native and mutant structures was carried out. We used the CLICK tool (http://cospi.iiserpune.ac.in/click/#) to calculate the Root Mean Square Deviation (RMSD) alignment.

For a significant match, the RMSD value should be above 0.15 and the Z-score above 2.5. The more important the match, the higher the Z-score [22]. In our case, all RMSD values are above 0.15, and z-score values are above 2.5, which means that our match is significant (Table 7).

Protein flexibility

Since proteins are highly dynamic macromolecules, many biological processes depend heavily on them. Numerous proteins must undergo conformational transformations in their folded states in order to carry out their intended functions, according to recent findings. Flexibility also enables the formation of complex structures through interactions with various partners, such as ligands via induced fit interactions, as well as other proteins or nucleic

Table 5. Homology modeling results.

Server	Template	Characteristics	3D Structure
SWISS MODEL	7f53.1	Identity:100% QMEAN :0.79 ERRAT=99.2509	
Phyre2	c6w25A	Identity: 97% QMEAN : -3.89 ERRAT=75.4717	
SPARKS-X	6gdgA	Identity: 26.5% QMEAN: -8.99 ERRAT=54.5455	
Robetta	6w25A	Identity: 54.14% QMEAN: -2.27 ERRAT=Errors	

acids[23]. The results of our prediction made by PredyFlexy, are shown in Table 8. The results showed that the T246N, G243R, and C196Y mutations increased residue flexibility while the C271R and D126Y mutations decreased it.

Dynamic cross-correlation matrix analysis using the DynOmics server

In order to better understand the interconnected communications between residues, a dynamic cross-correlation matrix was utilized. This tool enables the analysis of protein dynamics by assessing the fluctuation of both the native and mutant MC4R protein. Large fluctuations are represented in red colors, while small fluctuations are presented in blue. The results of the analysis revealed that the D298H, C271F, P260L, T246N, G243R, C196Y, and D90G variations led to a reduction in the degree of positive (red) and negative (blue) correlations, as compared to the wild-type protein MC4R.

Additionally, the results showed that the C271R and D126Y variations increased the degree of positive (red) and negative (blue) correlations found in the native MC4R, as compared to the wild type, while the E30K and P299L variants did not show any significant change (Table 9) (Figure 5).

Calculation of the cumulative score of the MCR4 deleterious missense variants

The combination of result from twelve algorithms shows that the following amino acid substitutions: E308K (rs375095163), P299L (rs52804924), D298H (rs760199460), C271R (rs1057517991), C271F (rs121913562), P260L (rs1435358988), T246N (rs1333658154), G243R (rs868309222), C196Y (rs1191554117), W174C (rs1159323398), Y157S (rs768916374), D126Y (rs768806551) and D90G (rs1598932350) were rated as the most plausible MC4R missense variants, with a cumulative score of 12 (Table S2).

Figure 3. Sequence conservation results.

Table 6. Evaluation of model quality generated by the Swiss Model.

NATIVE	Ramachandran Favoured: 98,21% QMEAN: 0,79 GMQE : 0,80 ERRAT :(Overall Quality Factor): 99,2509	T246N	Ramachandran Favoured: 97,86% QMEAN: 0,82 GMQE: 0,80 ERRAT :(Overall Quality Factor): 98,9011
E308K	Ramachandran Favoured: 98,21% QMEAN: 0,81 GMQE: 0,80 ERRAT :(Overall Quality Factor): 99,2509	**G243R**	Ramachandran Favoured: 97,86% QMEAN: 0,80 GMQE: 0,79 ERRAT :(Overall Quality Factor): 98,9011
P299L	Ramachandran Favoured: 98,57% QMEAN: O,80 GMQE :0,80 ERRAT :(Overall Quality Factor): 98,5019	**C196Y**	Ramachandran Favoured: 97,86% QMEAN: 0.81 GMQE : 0,80 ERRAT :(Overall Quality Factor): 98.9011
D298H	Ramachandran Favoured: 97,86% QMEAN: 0,80 GMQE: 0.79 ERRAT :(Overall Quality Factor): 98,5348	**W174C**	Ramachandran Favoured: 98,21% QMEAN: 0.81 GMQE: 0,80 ERRAT :(Overall Quality Factor): 99,2509
C271F	Ramachandran Favoured: 97.86% QMEAN: 0,80 GMQE: 0,79 ERRAT :(Overall Quality Factor): 98,9011	**Y157S**	Ramachandran Favoured: 98,21% QMEAN: 0,81 GMQE: 0,80 ERRAT :(Overall Quality Factor): 98,8764
C271R	Ramachandran Favoured: 98,21% QMEAN: 0,81 GMQE: 0,80 ERRAT :(Overall Quality Factor): 99,2509	**D126Y**	Ramachandran Favoured: 98,21% QMEAN: 0,81 GMQE: 0,80 ERRAT :(Overall Quality Factor): 98,8764
P260L	Ramachandran Favoured: 97,86% QMEAN: 0,81 GMQE: 0,79 ERRAT :(Overall Quality Factor): 98.9011	**D90G**	Ramachandran Favoured: 97,86% QMEAN: 0,81 GMQE: 0,80 ERRAT :(Overall Quality Factor): 98,5348

Table 7. Comparative modeling results.

Variant ID	mutation	RMSD	Z-score
rs375095163	E308K	0.39	20.64
rs52804924	P299L	0.27	20.90
rs760199460	D298H	0.55	20.04
rs121913562	C271F	0.55	20.04
rs1057517991	C271R	0.24	20.82
rs1435358988	P260L	0.52	20.19
rs1333658154	T246N	0.59	19.95
rs868309222	G243R	0.55	20.01
rs1191554117	C196Y	0.64	19.84
rs1159323398	W174C	0.55	20.01
rs768916374	Y157S	0.52	20.07
rs768806551	D126Y	0.40	20.61
rs1598932350	D90G	0.48	20.16

Discussion

Several factors play a role in the development of obesity, the most well-known of which are environmental factors (25-27), psychological and psychopathological factors (28), and genetic influences.

The twin studies provided tangible evidence that showed the involvement of genetics in the development of obesity. The studies had shown that even if the environmental factors and diet were different, two identical twins separated had the same risk of developing obesity (29).

Although obesity is due to complex interactions, the important and central role of genes in the development of the disease cannot be denied. Several MC4R SNPs have been identified as responsible for obesity in adults, such as rs12970134 (30,31), rs17782313 (32,33), rs571312 (35) , rs17700144 (36), rs571312 (37) and rs12970134 (37). Recently, studies have focused on the involvement of genetics in the development of early obesity in children, such as rs375095163 (38) , rs52804924 (39) , rs121913562 (40), and rs1057517991 (40).

The FTO gene (FAT MASS) was the first to spark researchers' interest in the development of childhood obesity (41). It was only in 1998 that the association between

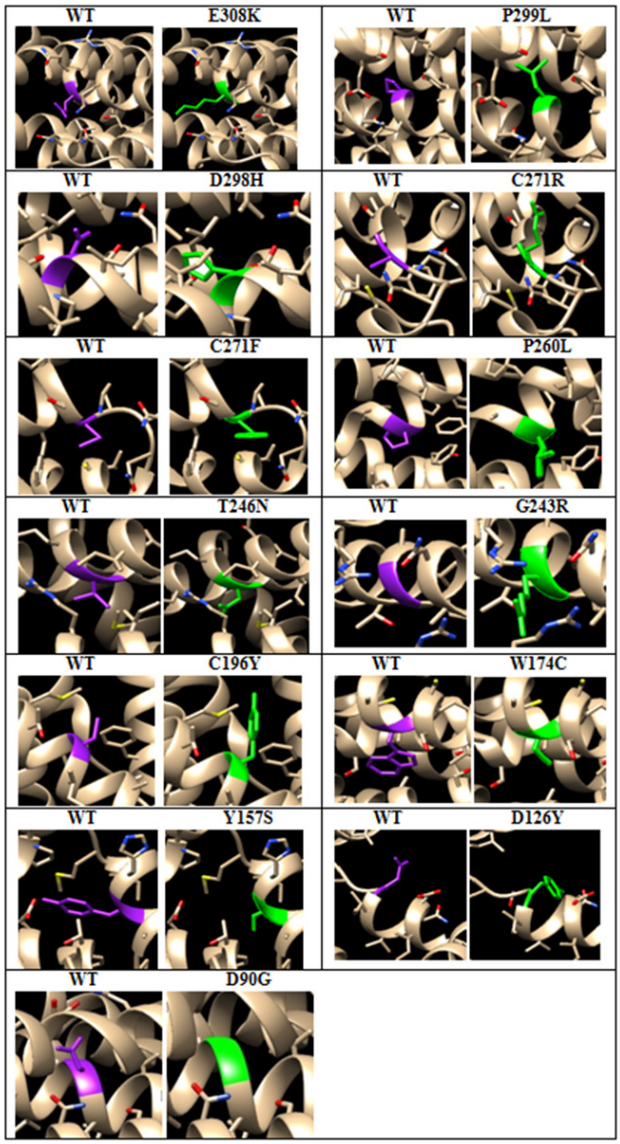


Figure 4. A close look at the mutation: The protein is shaded gray, and the side chains of the mutant and wild-type residues are visible at their intended locations in green and purple.

MC4R and childhood obesity was identified (42).

The MC4R gene found on18q21.32, encodes the MC4R protein, which is a G protein-coupled receptor that binds the α-Melanocyte- stimulating hormone. It regulates fullness and appetite and is mostly located in the hypothalamus (43). A number of studies, including genetic and pharmacological ones, have convincingly shown how crucial the MC4R is in controlling mice's food intake and energy homeostasis. Any nucleotide or amino acid level difference, polymorphism, or mutation can have a direct impact on how the body functions, leading to disease.

Our *in-silico* investigation presented the previous 13 mutations as very damaging substitutions that could cause disease. In the literature, only four nsSNPs (E308K, P299L, C271F, and C271R) (38-40,44) were considered to be associated with childhood obesity, and a mutation (W174C) was found to be associated with obesity in adults (45).

The E308K (rs375095163) mutation involves the substitution of glutamic acid (E) with Lysine (K) at position 308 of the MC4R protein. This mutation has been identified as damaging by seven tools and is located in a highly conserved and exposed region, as shown by the Consurf server. Its occurrence is predicted to decrease the stability of the MC4R protein. We used the ™-align tool to compare the mutant and native models and found a significant difference with an RMSD of 0,39. Based on these findings, we conclude that this mutation is likely to be deleterious and associated with disease. This conclusion is consistent with previous research linking the E308K mutation to childhood obesity. A case report of a 2 years old boy with early extreme obesity and heterozygous for the MC4R mutation (Glu308Lys) showed a sharp decrease in BMI after treatment with the stimulant methylphenidate. This study suggests a possible link between the melanocortinergic and dopamine systems and the sympathetic nervous system (38). Indeed, methylphenidate affects dopamine signaling, which seems to control eating behavior and weight. Decreasing the activity of dopamine by administering dopamine antagonists can result in an increase in food consumption and subsequent weight gain (46).

P299L (rs52804924) is a mutation that is predicted to be damaging by 7 tools. Its occurrence is likely to decrease the stability of the MC4R protein as it is located in a highly conserved and buried region, according to the Consurf server. Mutpred server has also shown that this mutation can cause several molecular changes, including a gain of the

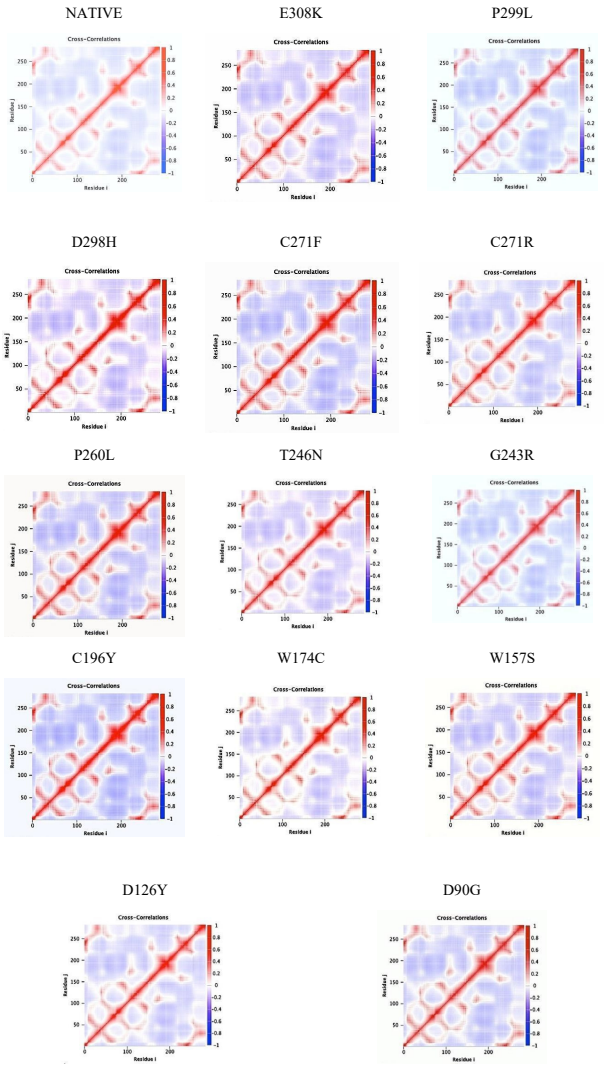


Figure 5. Dynamic analysis of the Cross-Correlation map of the MC4R protein native and mutants.

Table 8. Flexibility prediction results using the PredyFlexy server.

Variant	Confidence index	Predicted Flexibility Class	Effect of mutation on flexibility
E308	12	2	No effect on flexibility
E308K	12	2	
P299	8	1	No effect on flexibility
P299L	9	1	
D298	11	1	No effect on flexibility
D298H	11	1	
C271	7	1	No effect on flexibility
C271F	2	1	
C271	7	1	The mutation has made the residue rigid
C271R	17	0	
P260	8	1	No effect on flexibility
P260L	12	1	
T246	10	0	The mutation has made the residue flexible
T246N	15	2	
G243	13	0	The mutation has made the residue flexible
G243R	16	1	
C196	11	1	The mutation has increased the flexibility of the residue
C196Y	11	2	
W174	13	1	No effect on flexibility
W174C	13	1	
Y157	11	2	No effect on flexibility
Y157S	10	2	
D126	13	1	The mutation has made the residue rigid
D126Y	15	0	
D90	15	1	No effect on flexibility
D90G	13	1	

Legend: prediction class (0): rigid, (1): intermediate and (2): Flexible.

Table 9. The correlation between the predicted and observed fluctuations of both the native and mutant forms of the MC4R protein.

Mc4r structures	Native	E308K	P299L	D298H	C271F	C271R	P260L	T246N	G243R	C196Y	W174C	Y157S	D126Y	D90G
Correlation between observed and predicted fluctuations	0,76	0,76	0,76	0,62	0,63	0,87	0,64	0,54	0,63	0,53	0,76	0,87	0,87	0,72

helix, an altered ordered interface, a loss of an allosteric site at Y302, and altered metal binding, with a significant score of 0,839. The ™-align tool was used to compare the mutant model with the native model, and the result indicated a significant difference and non-superimposition between the two structures, with and RMSD of 0,27. Therefore, this mutation is considered to be deleterious and associated with diseases. This finding is consistent with a study that reported the association of the P299L mutation with childhood obesity, where three patients in the cohort were affected. Functional analysis of this mutant receptor also revealed that it alters the activation of the endogenous alpha-MSH agonist, where the MC4R alpha-MSH activation is completely removed from Pro299His (39).

C271F (rs121913562) is a mutation that has been predicted to be damaging by 7 tools. The occurrence of this mutation is likely to decrease the stability of the MC4R protein due to its location in a highly conserved and buried region, as reported by the Consurf server. The Mutpred server has also shown that the mutation can cause several molecular changes, such as altered transmembrane proteins, loss of helix, altered ordered interface, and gain of sulfation at Y268, with a significant result score of 0,927. Additionally, the ™-align tool was used to evaluate the superimposition of the mutant model on the native model, which revealed a significant difference and non-superim-

position between the two structures with an RMSD of 0.55. Therefore, we assume that this mutation is deleterious and may be associated with diseases. Notably, this mutation has been reported in ClinVar, which is a freely accessible archive of reports on the relationships between human variations and phenotypes (40).

C271R (rs1057517991) is a substitution of cysteine (C) with arginine (R) at position 271 that has been predicted to be damaging by 7 different tools. Its occurrence can lead to a decrease in the stability of the MC4R protein since it is situated in a highly conserved and buried area, as indicated by the Consurf server. Mutpred server analysis shows significant results (score=0.952) and indicates that this mutation can cause various molecular changes such as altered ordered transmembrane proteins, loss of helix, gain of stand, gain of loop, altered ordered interface, and loss of sulfatation at Y268. We also used the ™-align tool to compare the mutant model with the native model, and the result revealed a significant difference and non-superimposition between the two structures with an RMSD of 0.24. Therefore, we assume that this mutation is deleterious and associated with disease. A study conducted by Tarnow et al. supports this assumption by demonstrating a link between the C271R mutation and childhood obesity. The study showed that the C271R mutation can cause an inappropriate disulfide link between Cys-277 and Cys-279, which interferes with ligand binding and effective cell surface expression. The authors demonstrate that the absence of a Cys residue implicated in disulfide bonding promotes loss of function (47,44).

W174C (rs1159323398) results in the substitution of Tryptophan (W) with cysteine (C) at position 174. According to 7 prediction tools, this mutation is damaging and its occurrence is likely to decrease the stability of the MC4R protein, as it is located in a highly conserved and buried region, as revealed by the Consurf server. The Mutpred server also indicates that the mutation could cause several molecular changes, including gains of helix, stand, and loop, altered ordered interface, and loss of strand and transmembrane protein, with a significant result score of 0.946. Furthermore, using the ™-align tool, we found a significant difference and non-superimposition between the mutant model and the native model, with an RMSD of 0.55. Although there are no studies linking this mutation to childhood obesity, a study conducted on a large number of severely obese adults in southern Italy has shown its association with adult obesity (45).

For the other mutations (D298H, P260L, T246N, G243R, C196Y, Y157S, D126Y, and D90G), although they have proved to be very deleterious, they have not been found in the literature.

D298H (rs760199460) mutation refers to an amino acid substitution at position 298 where aspartic acid (D) is replaced by histidine (H). This mutation has been predicted to be damaging by 7 tools, and its occurrence is located in a highly conserved and buried region, as confirmed by the Consurf server. Furthermore, the Mutpred server has shown with a significant score of 0.874 that this mutation can generate various molecular-level changes, including an altered ordered interface, and altered metal binding. Additionally, we have used the ™-align tool to analyze whether the mutant model is superimposed on the native model. The comparison has revealed a significant difference and non-superimposition between the two structures, with an RMSD of 0.55.

P260L (rs1435358988) is a substitution of proline (P) to leucine (L) at position 260, which has been predicted by seven tools to be damaging. The occurrence of this mutation will result in decreased stability of the MC4R protein due to its placement in a highly conserved and buried region, as evidenced by the Consurf server. Mutpred server has revealed a significant results score of 0.843, indicating that the mutation can cause several molecular changes, such as the loss of strand, gain of the helix, altered ordered interface, an altered transmembrane protein, and altered metal binding. To assess the impact of this mutation further, we utilized the ™-align tool to compare the mutant model with the native model, which revealed a significant difference and non-superimposition between the two structures with an RMSD of 0.52. No study to date has associated this mutation with childhood obesity or adult obesity.

T246N (rs1333658154) is a substitution of threonine (T) to aspargine (N) at position 246, which has been predicted to be damaging by 7 tools. This mutation is located in a highly conserved and buried region, as indicated by the Consurf server, and its occurrence is expected to decrease the stability of the protein. The Mutpred server score of 0.885 suggests that the mutation can cause several molecular changes, including an altered transmembrane protein, the gain of a helix, the loss of a strand, the loss of a catalytic site at K242, the loss of methylation at K242, and the gain of N-linked glycosylation at T246. To assess the structural impact of the mutation, we compared the mutant model with the native model using the ™-align tool, which showed a significant difference and non-superimposition between the two structures with an RMSD of 0.59.

G243R (rs868309222) is a substitution of glycine (G) to arginine (R) at position 243. Like T246, this mutation has been predicted to be damaging by 7 tools, and its occurrence is expected to decrease the stability of the protein. However, it is located in a highly conserved and exposed region, as indicated by the Consurf server. The Mutpred server score of 0.956 suggests that the mutation can cause several molecular changes, including the gain of a helix, an altered transmembrane protein, an altered ordered interface, the loss of a strand, the loss of a catalytic site at K242, the loss of methylation at K242, and the loss of GPI-anchor amidation at N240. The comparison of the mutant model with the native model using the ™-align tool showed a significant difference and non-superimposition between the two structures with an RMSD of 0.55.

C196Y (rs1191554117) is a substitution of cysteine (C) for tyrosine (Y) at position 196. Similar to the other two mutations, it has been predicted to be damaging by 7 tools and is located in a highly conserved and buried region, as indicated by the Consurf server. The Mutpred server score of 0.914 suggests that the mutation can cause several molecular changes, including altered transmembrane proteins and altered ordered interfaces. The comparison of the mutant model with the native model using the ™-align tool showed a significant difference and non-superimposition between the two structures with an RMSD of 0.64.

Y157S (rs768916374) is an amino acid substitution where Tyrosine (Y) is replaced by Serine (S) at position 157. This substitution has been predicted to be damaging by 7 tools. Its occurrence is expected to reduce the stability of the MC4R protein, which is situated in a highly conser-

ved and buried region, as indicated by the Consurf server. The Mutpred server has shown that the mutation can lead to several molecular-level changes with a significant score of 0.915, including an altered ordered interface, altered transmembrane protein, loss of an allosteric site at Y157, altered metal binding, and altered DNA binding. We utilized the ™-align tool to compare the mutant and native models, revealing a significant difference and non-superimposition between the two structures with an RMSD of 0.52.

D126Y (rs768806551) is an amino acid substitution where aspartic acid (D) is replaced by glycine (Y) at position 126. This mutation has been predicted to be damaging by 7 tools. Its occurrence is expected to decrease the stability of the MC4R protein, which is located in a highly conserved and exposed region, as indicated by the Consurf server. The Mutpred server has shown that the mutation can lead to several molecular-level changes with a significant score of 0.952, including an altered transmembrane protein, an altered layered ordered interface, an altered transmembrane protein, a gain of the strand, a loss of helix, altered metal binding, and a gain of sulfation at D126. We utilized the ™-align tool to compare the mutant and native models, revealing a significant difference and non-superimposition between the two structures with an RMSD of 0.40.

D90G (rs1598932350) is an amino acid substitution where aspartic acid (D) is replaced by glycine (G) at position 90. This mutation has also been predicted to be damaging by 7 tools. Its occurrence is expected to decrease the stability of the MC4R protein, which is located in a highly conserved and exposed region, as indicated by the Consurf server. The Mutpred server has shown that the mutation can lead to several molecular-level changes with a significant score of 0.918, including loss of the helix and altered transmembrane proteins. We utilized the ™-align tool to compare the mutant and native models, revealing a significant difference and non-superimposition between the two structures with an RMSD of 0.48.

On the other hand, some mutations have been proven to have an association with the onset of childhood obesity but that have not been shown to be deleterious in our *in-silico* analysis. We can quote:

T150I (rs766665118) is an amino acid substitution at position 150 from threonine (T) to isoleucine (I). A study in Spanish children with abdominal obesity showed that Thr150Ile is a mutation leading to a reduced function of the *MC4R* exerts a major gene effect on extreme childhood obesity (48).

S136F (rs1380965800) is an amino acid substitution at position 136 from serine (S) to phenylalanine(F).A 2.3-year-old girl with severe obesity had the mutation. (BMI: 33.2 kg/m(2), >99th centile). According to the study, S136F co-segregated for three generations with the obese phenotype. While the mutant receptor's cell surface expression was barely changed in comparison to the wild-type receptor, in vitro functional investigations showed that the mutant receptor completely lost its signal transduction activity (49).

Out of the sixteen nsSNPs found in the MC4R protein, thirteen were considered to have the greatest potential for pathogenicity and impact on protein stability. These specific nsSNPs were selected for further analysis. Our findings indicate that each of the thirteen substitutions produced a significant degree of pathogenicity, as confirmed by all the tools and reflected in the cumulative score (see Table S2). While some of these nsSNPs have not yet been experimentally linked to the disease, our study provides valuable information for teams researching this condition. However, we acknowledge that there are limitations to our study.

This in silico study, has certain limitations to be taken into account. Indeed, they are based on models and algorithms and require experimental validation to confirm the results. Moreover, the results of in silico studies provide predictions on the functional impact of genetic variations, but their interpretation requires in-depth analysis and context with other clinical and experimental data. Thus, in silico studies are an important step, but they must be complemented by experimental studies to validate the results and fully understand their impact on complex diseases.

In terms of future research directions, experimental validation through in vitro and in vivo studies is necessary to confirm the functional impact of the identified missense variations. Investigating gene-gene interactions related to obesity and considering the unique genetic profiles of diverse populations would provide a deeper understanding of the genetic factors contributing to childhood obesity. Longitudinal studies following individuals with specific genetic variations from childhood to adulthood would shed light on the long-term implications on obesity risk and associated health outcomes. These research efforts would contribute to personalized approaches for obesity prevention and management based on an individual's genetic profile.

Conclusion

The present study identified 13 nsSNPs in the MC4R gene through in silico-analysis of missense variants (E308K, P299L, D298H, C271R, P260L, T246N, C196Y, W174C, Y157S, D126Y, and D90G). These nsSNPs were found to be the most deleterious among all known and reported MC4R gene nsSNPs. By pointing out mutations that potentially impact the function of the MC4R protein, we can identify subpopulations of individuals who may benefit from a precision medicine strategy that aims to pharmacologically compensate for genetic deficits in the MC4R pathway.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors are thankful to everyone who contributed to this work for providing encouragement and facilities.

Supplementary file

Supplemental file 1: Table S2. Prediction of molecular change of substitutions by MutPred.
Supplemental file2: Table S10. Result of selected pathogenic nsSNPs according to all prediction tools used.

References

1. Goran, M. I. Metabolic Precursors and Effects of Obesity in Children: A Decade of Progress, 1990-1999. Am J Clin Nutr.2001, 73 (2), 158–171. https://doi.org/10.1093/ajcn/73.2.158.
2. Stunkard, A. J.; Foch, T. T.; Hrubec, Z. A Twin Study of Human Obesity. JAMA, 1986, 256 (1), 51–54.

3. Montagne, L. Génétique de l'obésité de l'enfant. Médecine humaine et pathologie, Université du droit et de la santé : LILLE 2, 2017.
4. da Fonseca, A. C. P.; Mastronardi, C.; Johar, A.; Arcos-Burgos, M.; Paz-Filho, G. Genetics of Non-Syndromic Childhood Obesity and the Use of High-Throughput DNA Sequencing Technologies. J DC. 2017, 31 (10), 1549–1561. https://doi.org/10.1016/j.jdiacomp.2017.04.026.
5. Tao, Y.-X. The Melanocortin-4 Receptor: Physiology, Pharmacology, and Pathophysiology. Endocr Rev.2010, 31 (4), 506–543. https://doi.org/10.1210/er.2009-0037.
6. Zlatohlávek, L.; Hubáček, J. A.; Vrablík, M.; Pejšová, H.; Lánská, V.; Češka, R. The Impact of Physical Activity and Dietary Measures on the Biochemical and Anthropometric Parameters in Obese Children. Is There Any Genetic Predisposition? Cent Eur J Public Health.2015, 23 Suppl, S62-66. https://doi.org/10.21101/cejph.a4191.
7. Tao, Y.-X. Molecular Mechanisms of the Neural Melanocortin Receptor Dysfunction in Severe Early Onset Obesity. Mol Cell Endocrinol.2005, 239 (1–2), 1–14. https://doi.org/10.1016/j.mce.2005.04.012.
8. Johnson, A. D. Single-Nucleotide Polymorphism Bioinformatics: A Comprehensive Review of Resources. Circ Cardiovasc Genet.2009, 2 (5), 530–536. https://doi.org/10.1161/CIRCGENETICS.109.872010.
9. C Ng, P.; Henikoff, S. SIFT: Predicting amino acid changes that affect protein function - PubMed. Nucleic Acids Res.2003, 31, 3812. https://doi.org/10.1093/nar/gkg509.
10. Xi, T.; Jones, I. M.; Mohrenweiser, H. W. Population Screenings Are Predicted to Impact Protein Function. 2004, 26.
11. Hecht, M.; Bromberg, Y.; Rost, B. News from the Protein Mutability Landscape. J Mol Biol.2013, 425 (21), 3937–3948. https://doi.org/10.1016/j.jmb.2013.07.028.
12. Choi, Y.; Sims, G. E.; Murphy, S.; Miller, J. R.; Chan, A. P. Predicting the Functional Effect of Amino Acid Substitutions and Indels. PLOS ONE.2012, 7 (10), e46688. https://doi.org/10.1371/journal.pone.0046688.
13. Calabrese, R.; Capriotti, E.; Fariselli, P.; Martelli, P. L.; Casadio, R. Functional Annotations Improve the Predictive Score of Human Disease-Related Mutations in Proteins. Hum Mutat.2009, 30 (8), 1237–1244. https://doi.org/10.1002/humu.21047.
14. Johnson, A. D.; Handsaker, R. E.; Pulit, S. L.; Nizzari, M. M.; O'Donnell, C. J.; de Bakker, P. I. W. SNAP: A Web-Based Tool for Identification and Annotation of Proxy SNPs Using HapMap. Bioinformatics.2008, 24 (24), 2938–2939. https://doi.org/10.1093/bioinformatics/btn564.
15. Cheng, J.; Randall, A.; Baldi, P. Prediction of Protein Stability Changes for Single-Site Mutations Using Support Vector Machines. Proteins.2006, 62 (4), 1125–1132. https://doi.org/10.1002/prot.20810.
16. Capriotti, E.; Fariselli, P.; Calabrese, R.; Casadio, R. Predicting Protein Stability Changes from Sequences Using Support Vector Machines. Bioinformatics.2005, 21 (suppl_2), ii54–ii58. https://doi.org/10.1093/bioinformatics/bti1109.
17. Armon, A.; Graur, D.; Ben-Tal, N. ConSurf: An Algorithmic Tool for the Identification of Functional Regions in Proteins by Surface Mapping of Phylogenetic Information. Journal of Molecular Biology.2001, 307 (1), 447–463. https://doi.org/10.1006/jmbi.2000.4474.
18. Chothia, C.; Lesk, A. M. Canonical Structures for the Hypervariable Regions of Immunoglobulins. J. Mol. Biol.1987, 196 (4), 901–917. https://doi.org/10.1016/0022-2836(87)90412-8.
19. Benkert, P.; Tosatto, S. C. E.; Schomburg, D. QMEAN: A Comprehensive Scoring Function for Model Quality Assessment. Proteins.2008, 71 (1), 261–277. https://doi.org/10.1002/prot.21715.
20. Colovos, C.; Yeates, T. O. Verification of Protein Structures: Patterns of Nonbonded Atomic Interactions. Protein Sci.1993, 2 (9), 1511–1519. https://doi.org/10.1002/pro.5560020916.
21. Couch, G. S.; Hendrix, D. K.; Ferrin, T. E. Nucleic Acid Visualization with UCSF Chimera. Nucleic Acids Res.2006, 34 (4), e29–e29. https://doi.org/10.1093/nar/gnj031.
22. Nguyen, M. N.; Tan, K. P.; Madhusudhan, M. S. CLICK—Topology-Independent Comparison of Biomolecular 3D Structures. Nucleic Acids Res.2011, 39 (suppl_2), W24–W28. https://doi.org/10.1093/nar/gkr393.
23. de Brevern, A. G.; Bornot, A.; Craveur, P.; Etchebest, C.; Gelly, J.-C. PredyFlexy: Flexibility and Local Structure Prediction from Sequence. Nucleic Acids Res.2012, 40 (W1), W317–W322. https://doi.org/10.1093/nar/gks482.
24. H, L.; Yy, C.; Jy, L.; I, B.; Lw, Y. DynOmics: dynamics of structural proteome and beyond. Nucleic acids res.2017, 45 (W1). https://doi.org/10.1093/nar/gkx385.
25. Dennison, B. A.; Erb, T. A.; Jenkins, P. L. Television Viewing and Television in Bedroom Associated with Overweight Risk among Low-Income Preschool Children. Pediatrics.2002, 109 (6), 1028–1035. https://doi.org/10.1542/peds.109.6.1028.
26. Must, A.; Tybor, D. J. Physical Activity and Sedentary Behavior: A Review of Longitudinal Studies of Weight and Adiposity in Youth. Int J Obes (Lond).2005, 29 Suppl 2, S84-96. https://doi.org/10.1038/sj.ijo.0803064.
27. Marshall, S. J.; Biddle, S. J. H.; Gorely, T.; Cameron, N.; Murdey, I. Relationships between Media Use, Body Fatness and Physical Activity in Children and Youth: A Meta-Analysis. Int J ObesRelatMetab Disord.2004, 28 (10), 1238–1246. https://doi.org/10.1038/sj.ijo.0802706.
28. Blaine, B. Does Depression Cause Obesity?: A Meta-Analysis of Longitudinal Studies of Depression and Weight Control. J Health Psychol.2008, 13 (8), 1190–1197. https://doi.org/10.1177/1359105308095977.
29. Stunkard, A. J.; Foch, T. T.; Hrubec, Z. A Twin Study of Human Obesity. JAMA.1986, 256 (1), 51–54.
30. Thorleifsson, G.; Walters, G. B.; Gudbjartsson, D. F.; Steinthorsdottir, V.; Sulem, P.; Helgadottir, A.; Styrkarsdottir, U.; Gretarsdottir, S.; Thorlacius, S.; Jonsdottir, I.; et al. Genome-Wide Association Yields New Sequence Variants at Seven Loci That Associate with Measures of Obesity. Nat Genet.2009, 41 (1), 18–24. https://doi.org/10.1038/ng.274.
31. Lo, M.-T.; Hinds, D. A.; Tung, J. Y.; Franz, C.; Fan, C.-C.; Wang, Y.; Smeland, O. B.; Schork, A.; Holland, D.; Kauppi, K.; et al. Genome-Wide Analyses for Personality Traits Identify Six Genomic Loci and Show Correlations with Psychiatric Disorders. Nat Genet.2017, 49 (1), 152–156. https://doi.org/10.1038/ng.3736.
32. Loos, R. J. F.; Lindgren, C. M.; Li, S.; Wheeler, E.; Zhao, J. H.; Prokopenko, I.; Inouye, M.; Freathy, R. M.; Attwood, A. P.; Beckmann, J. S.; et al. Common Variants near MC4R Are Associated with Fat Mass, Weight and Risk of Obesity. Nat Genet.2008, 40 (6), 768–775. https://doi.org/10.1038/ng.140.
33. Willer, C. J.; Speliotes, E. K.; Loos, R. J. F.; Li, S.; Lindgren, C. M.; Heid, I. M.; Berndt, S. I.; Elliott, A. L.; Jackson, A. U.; Lamina, C.; et al. Six New Loci Associated with Body Mass Index Highlight a Neuronal Influence on Body Weight Regulation. Nat Genet.2009, 41 (1), 25–34. https://doi.org/10.1038/ng.287.
34. Meyre, D.; Delplanque, J.; Chèvre, J.-C.; Lecoeur, C.; Lobbens, S.; Gallina, S.; Durand, E.; Vatin, V.; Degraeve, F.; Proença, C.; et al. Genome-Wide Association Study for Early-Onset and Morbid Adult Obesity Identifies Three New Risk Loci in European Populations. Nat Genet.2009, 41 (2), 157–159. https://doi.org/10.1038/ng.301.

35. Speliotes, E. K.; Willer, C. J.; Berndt, S. I.; Monda, K. L.; Thorleifsson, G.; Jackson, A. U.; Allen, H. L.; Lindgren, C. M.; Luan, J.; Mägi, R.; et al. Association Analyses of 249,796 Individuals Reveal 18 New Loci Associated with Body Mass Index. Nat Genet.2010, 42 (11), 937–948. https://doi.org/10.1038/ng.686.
36. Scherag, A.; Dina, C.; Hinney, A.; Vatin, V.; Scherag, S.; Vogel, C. I. G.; Müller, T. D.; Grallert, H.; Wichmann, H.-E.; Balkau, B.; et al. Two New Loci for Body-Weight Regulation Identified in a Joint Analysis of Genome-Wide Association Studies for Early-Onset Extreme Obesity in French and German Study Groups. PLoS Genet.2010, 6 (4), e1000916. https://doi.org/10.1371/journal.pgen.1000916.
37. Gao, L.; Wang, L.; Yang, H.; Pan, H.; Gong, F.; Zhu, H. MC4R Single Nucleotide Polymorphisms Were Associated with Metabolically Healthy and Unhealthy Obesity in Chinese Northern Han Populations. Int J Endocrinol.2019, 2019, 4328909. https://doi.org/10.1155/2019/4328909.
38. Albayrak, Ö.; Albrecht, B.; Scherag, S.; Barth, N.; Hinney, A.; Hebebrand, J. Successful Methylphenidate Treatment of Early Onset Extreme Obesity in a Child with a Melanocortin-4 Receptor Gene Mutation and Attention Deficit/Hyperactivity Disorder. European Journal of Pharmacology.2011, 660 (1), 165–170. https://doi.org/10.1016/j.ejphar.2010.12.023.
39. Lubrano-Berthelier, C.; Durand, E.; Dubern, B.; Shapiro, A.; Dazin, P.; Weill, J.; Ferron, C.; Froguel, P.; Vaisse, C. Intracellular Retention Is a Common Characteristic of Childhood Obesity-Associated MC4R Mutations. Human Molecular Genetics.2003, 12 (2), 145–153. https://doi.org/10.1093/hmg/ddg016.
40. Landrum, M. J.; Lee, J. M.; Benson, M.; Brown, G. R.; Chao, C.; Chitipiralla, S.; Gu, B.; Hart, J.; Hoffman, D.; Jang, W.; et al. ClinVar: Improving Access to Variant Interpretations and Supporting Evidence. Nucleic Acids Res.2018,46 (Database issue), D1062–D1067. https://doi.org/10.1093/nar/gkx1153.
41. Frayling, T. M.; Timpson, N. J.; Weedon, M. N.; Zeggini, E.; Freathy, R. M.; Lindgren, C. M.; Perry, J. R. B.; Elliott, K. S.; Lango, H.; Rayner, N. W.; et al. A Common Variant in the FTO Gene Is Associated with Body Mass Index and Predisposes to Childhood and Adult Obesity. Science. 2007, 316 (5826), 889–894. https://doi.org/10.1126/science.1141634.
42. Yeo, G. S.; Farooqi, I. S.; Aminian, S.; Halsall, D. J.; Stanhope, R. G.; O'Rahilly, S. A Frameshift Mutation in MC4R Associated with Dominantly Inherited Human Obesity. Nat Genet.1998, 20 (2), 111–112. https://doi.org/10.1038/2404.
43. Dubern, B.; Clément, K.; Pelloux, V.; Froguel, P.; Girardet, J.-P.; Guy-Grand, B.; Tounian, P. Mutational Analysis of Melanocortin-4 Receptor, Agouti-Related Protein, and α-Melanocyte-Stimulating Hormone Genes in Severely Obese Children. J Pediatr.2001, 139 (2), 204–209. https://doi.org/10.1067/mpd.2001.116284.
44. Farooqi, I. S.; Keogh, J. M.; Yeo, G. S. H.; Lank, E. J.; Cheetham, T.; O'Rahilly, S. Clinical Spectrum of Obesity and Mutations in the Melanocortin 4 Receptor Gene. N Engl J Med.2003, 348 (12), 1085–1095. https://doi.org/10.1056/NEJMoa022050.
45. Buono, P.; Pasanisi, F.; Nardelli, C.; Ieno, L.; Capone, S.; Liguori, R.; Finelli, C.; Oriani, G.; Contaldo, F.; Sacchetti, L. Six Novel Mutations in the Proopiomelanocortin and Melanocortin Receptor 4 Genes in Severely Obese Adults Living in Southern Italy. Clin. Chem.2005, 51 (8), 1358–1364. https://doi.org/10.1373/clinchem.2005.047886.
46. Baptista, T. Body Weight Gain Induced by Antipsychotic Drugs: Mechanisms and Management. Acta Psychiatr Scand.1999, 100 (1), 3–16. https://doi.org/10.1111/j.1600-0447.1999.tb10908.x.
47. Tarnow, P.; Schoneberg, T.; Krude, H.; Gruters, A.; Biebermann, H. Mutationally Induced Disulfide Bond Formation within the Third Extracellular Loop Causes Melanocortin 4 Receptor Inactivation in Patients with Obesity. J Biol Chem.2003, 278 (49), 48666–48673. https://doi.org/10.1074/jbc.M309941200.
48. Morell-Azanza, L.; Ojeda-Rodríguez, A.; Giuranna, J.; Azcona-SanJulián, M. C.; Hebebrand, J.; Marti, A.; Hinney, A. Melanocortin-4 Receptor and Lipocalin 2 Gene Variants in Spanish Children with Abdominal Obesity: Effects on BMI-SDS after a Lifestyle Intervention. Nutrients.2019, 11 (5). https://doi.org/10.3390/nu11050960.
49. Rettenbacher, E.; Tarnow, P.; Brumm, H.; Prayer, D.; Wermter, A.-K.; Hebebrand, J.; Biebermann, H.; Hinney, A.; Widhalm, K. A Novel Non-Synonymous Mutation in the Melanocortin-4 Receptor Gene (MC4R) in a 2-Year-Old Austrian Girl with Extreme Obesity. Exp Clin Endocrinol Diabetes.2007, 115 (1), 7–12. https://doi.org/10.1055/s-2007-949150.