OPEN ACCESS

Haya: The Saudi Journal of Life Sciences
Abbreviated Key Title: Haya Saudi J Life Sci
ISSN 2415-623X (Print) | ISSN 2415-6221 (Online)
Scholars Middle East Publishers, Dubai, United Arab Emirates
Journal homepage: https://saudijournals.com

Original Research Article

Exploring the Ultra-Rare Truncating Protein Variant Missense Mutation and Regulatory SNPs of the Human *PRDM16* Using in *Silico* Approach

Amrit Sudershan[1,2], Shikha Bharti[3], Meenakshi Bhagat[4], Isha Sharma[4], Rakesh K Panjalyia[4], Parvinder Kumar[1,4*]

[1]Institute of Human Genetics, University of Jammu, Jammu, Jammu & Kashmir, 180006
[2]Department of Human Genetics, Sri Pratap College, Cluster University of Srinagar, Kashmir, Jammu & Kashmir, India-190001
[3]Department of Zoology, Lovely Professional University, Punjab, 144411
[4]Department of Zoology, University of Jammu, Jammu, Jammu & Kashmir, 180006

DOI: 10.36348/sjls.2023.v08i10.003 | **Received:** 18.09.2023 | **Accepted:** 23.10.2023 | **Published:** 16.11.2023

***Corresponding author:** Parvinder Kumar
Institute of Human Genetics, University of Jammu, Jammu, Jammu & Kashmir, 180006

Abstract

***Background*:** Genomics is one of the disciplines of modern medicine that focuses on identifying causative genes and their related variations that may have an impact on complex disorders. Candidate gene association studies are critical for determining the genetic relationship of genomic variations with complicated illnesses. ***Aim*:** The goal of this study is to anticipate the likely relationship of PRDM16 gene variations with negative effects on structural and functional features using online computational tools. ***Methodology*:** An insilico approach was utilized to find out the rare variant in the PRDM16 gene. ***Result*:** We found eight missense variants including rs572205989, rs201814961, rs572178955, rs182452331, rs551202646, rs554705536, rs184929979 and rs573567598that could play a role in the development of disease. ***Discussion & conclusion*:** This methodology can be used in future genomes and association studies, but it must be tested in a model organism and cell culture. This research could be useful in personalized therapy and could lead to the discovery of new therapeutic markers for a variety of disorders.

Keywords: Insilico tools, Sequencing, SNP, Genomics.

1. INTRODUCTION

Various human disorders do not exhibit Mendelian inheritance patterns (Sladek *et al*., 2007) therefore called complex disorders and are caused by several genetic variants in the genes (polygenic inheritance) with modest effect together (Boyle *et al*., 2017) in combination with environmental factors (multifactorial inheritance) thereby posing a significant risk for the particular disorder (Bhatti *et al*., 2006). Molecular epidemiology research studies have identified a plethora of genetic variations and their associated heritable phenotype (Timpson *et al*., 2018). However, finding these genetic variations for the specific phenotypes is difficult and requires multiple testing of hundreds or even thousands of Single Nucleotide Polymorphisms (SNPs) in the candidate gene. The candidate gene studies approach has been at the forefront of uncovering the correlation between genetic variants complex diseases (Patnala *et al*., 2013). In recent years high- throughput techniques such as Genome-Wide Association Studies (GWAS) generated several common SNPs in candidate genes (Manolio, 2010; Tam *et al*., 2019) among which only a small fraction is truly relevant to the phenotype of interest (Tranchevent *et al*., 2016).

Migraine, a polygenic, dysautonomic, complex neurological disorder with vascular dysfunction as a consequence is characterized by the "lower threshold of neuronal hyper-excitability" called "the migrainous brain" (Kogelman *et al*., 2020; Sudershan *et al*., 2022; Sudershan *et al*., 2023). Pathogenesis of migraine have been explored since following decades (Sudershan *et al*., 2023; Leao, 1944; Bolay *et al*., 2002) and also showed varied prevalence in different population (Andreou& Edvinsson, 2019; Sudershan *et al*., 2023). A meta-analysis of 22 GWA studies conducted by Gormley *et al*., which was different from the previous GWA studies (Anttila *et al*., 2010' Esserlind *et al*., 2013; Ligthart *et al*., 2011; Anttila *et al*., 2013; Freilinger *et al*., 2012;) identified 38 genomic loci harboring 44 independent susceptibility markers for the prevalent forms of migraine (Gormley *et al*., 2016). A replication study of GWAS conducted by Ghosh *et al*., and Kaur *et al*., showed that rs2651899 (*PRDM16*) is a potential genetic marker for migraine susceptibility in MWA (Migraine

Without Aura) and female subgroup at both genotypic and allelic levels in the North Indian population (Ghosh *et al*., 2013; Kaur *et al*., 2019).

PR domain containing 16 (PRDM16) (NCBI ID: 63976) is located on chromosome 1p36.32 and consists of 17 exons with a protein length of 1276 amino acids. The protein binds to the double-stranded (ds) or single-stranded DNA and is known to play a function as a transcriptional regulator. It is also shown to display histone methyltransferases (HMT) activity (https://www.omim.org/entry/605557).

2. MATERIAL AND METHOD

The number of bioinformatics tools has been built dramatically over the two past decades. Literature was surfed on different online libraries including PubMed, and Google Scholar, using searches "*in-silico* tools for SNPs", "online tools used to find SNPs", "freely available tools for prioritizing SNPs, bioinformatics tools for prioritizing, etc.

2.1. Data Mining

The GnomAD genome browser (https://gnomad.broadinstitute.org/), UCSC genome browser (https://genome.ucsc.edu/), and NCBI Variation Viewer (https://www.ncbi.nlm.nih.gov/variation/view/) were used to obtain SNPs for the PRDM16 gene. The schematic depiction for selecting SNPs and using computational tools for PRDM16 in silico analysis is shown (Figure 1).

2.2. Bioinformatics Data Analysis

VEP Ensembl

The Ensembl Variant Effect Predictor (VEP Ensembl) is a powerful web-based tool used for the analysis, annotation, and prioritization of genomic variants in the coding and non-coding regions which helps in simplifying and accelerates the variant interpretation and it also give information related to the impact of a variant on a transcript or a protein and predicts biophysical consequences of variants using various prediction tools including SIFT, Polyphene 2, ConDel, etc. Many web-based methods were applied in this work to prepare the functional impact and pathogenic character of nsSNPs. All tools that were used in the current study were from Variant Effect Predictor Ensemble (VEP) (https://asia.ensembl.org/info/docs/tools/vep/index.html) and Variant Annotation Integrator (VAI) UCSC genome browser **(**https://genome.ucsc.edu/tools/) according to their default setting.

SIFT

SIFT (Sorting Intolerant from Tolerant), (https://sift.bii.a-star.edu.sg/) predicts whether an amino acid substitution causes a deleterious effect based on the sequence homology and the physical properties of amino acid. A Missense variant is predicted to be deleterious when the SIFT score is < 0.05, whereas ≥ 0.05 indicates that a variant is benign (Henikoff *et al*., 2003).

Polyphene2

PolyPhene2 (Polymorphism Phenotyping v2,http://genetics.bwh.harvard.edu/pph2/), an automatic web-based free tool that predicts the possible impact of amino acid substitutions on the stability and function of human proteins using structural and comparative evolutionary considerations. Different mutations are categorized as "possibly damaging", "probably damaging" and "Benign (0.0)". Variants with scores in the range of 0.0 to 0.15 are predicted to be benign, 0.15 to 0.85 are possibly damaging, and 0.85 to 1.0 are predicted to be damaging (Adzhubei *et al*., 2013).

Mutation Taster 2

Mutation Taster 2 (http://www.mutationtaster.org/) is designed to predict the functional consequences of genetic variants and evaluate the pathogenic potential of DNA sequence alterations that were downloaded from the genome browser. MutationTaster2 includes all publicly available single-nucleotide polymorphisms (SNPs) and indels from the 1000 Genomes Project2 (hereafter referred to as 1000G) as well as known disease variants from ClinVar3 and HGMD Public4 (Schwarz *et al.*, 2014). Signs used are Disease-causing SNP is denoted by A, disease-causing as D, polymorphism as N, and polymorphism as P (Schwarz *et al*., 2014).

Mutation Assessor

Mutation Assessor (http://mutationassessor.org/r3/) is a helpful prediction tool used to find the functional impact of amino-acid substitutions in proteins, such as missense polymorphisms. The assessment of the functional impact of amino-acid substitutions in proteins was based on the evolutionary conservation of the affected amino acids in protein homologs (Reva *et al*., 2011). Results are retrieved in high or medium; predicted functional; low or neutral; and predicted non-functional (Gnad *et al*., 2013).

Likelihood Ratio Test (LRT)

Likelihood Ratio Test (LRT) information from comparative genomics is taken by the LRT tool thus helping to identify deleterious mutations that disrupt highly conserved amino acids within protein-coding sequences (deleterious=D, Neutral=N, and Unknown=U) (Chun *et al*., 2009).

VEST

Variant Effect Scoring Tool (VEST) (Carter *et al*., 2013) is a new method for prioritizing missense mutations that alter protein activity.

2.3. Protein Stability Change

I-Mutant

I-mutant (http://gpcr2.biocomp.unibo.it/cgi/predictors/I-Mutant3.0/I-Mutant3.0.cgi) is a free web-based tool,

which is used to predict the stability of folded protein which changes upon single point mutations. Predictions were performed using protein sequence (http://gpcr2.biocomp.unibo.it/cgi/predictors/I-Mutant3.0/I-Mutant3.0.cgi) (Capriotti *et al*., 2005).

2.4. Regulatory SNPs

Mutations in gene promoter/regulatory regions represent an important class of lesions causing human genetic disease. Such mutations are associated with either increases or decreases in transcriptional activity and are called functional polymorphisms. These polymorphisms may be trans-acting or cis-acting, a polymorphism in which one gene affects the expression of another gene or the same gene respectively (Cooper, 1992).

Non-coding variants were downloaded from the NCBI dbSNP database and about 108659 total variants were obtained. After filtering we found 7321 number of Non-coding variants and 52 polymorphisms having MAF between 0.05 to 0.10 (5% to 10%). To filter and prioritize these 52 genomic variants hosting functional SNPs, the Regulome database was used and 44 variants were obtained. Lower scores indicate increasing evidence for a variant to be located in a functional region (Boyle *et al*., 2012) (table 4).

2.5. Gene-To-Gene Interaction Using String V11

String v11 (https://string-db.org/cgi/input/), is a potential protein-protein interaction tool that collects data from several online databases (Szklarczyk *et al*., 2019). It is very important to find out the most important wiring connection in the gene regulatory network that is used to understand the disease's development (Wray *et al*., 2018) (Figure 6) because any change in the peripheral gene will ultimately affect on the regulation of the core gene.

3. RESULTS

Identification of Nssnps in PRDM16 Gene and Functional Characterization of nsSNPs

By digging out the GnomAD browser, UCSC genome browser, and dbSNPs NCBI we found different numbers of missense variants. The total number of 1627, 966, and 839 was found using UCSC genome browser, dbSNPs NCBI, and GnomAD database respectively (Figure 4). All these variants were then filtered using VEP Ensembl using criteria “excluding common variants” and obtained 152, 195, and 113 (UCSC genome browser, dbSNPs NCBI, and GnomAD database respectively missense variants (Figure 4). These missense variants were then selected for the further filtering process.

Selection of Deleterious nsSNPs

For the selection of deleterious nsSNPs, two different criteria were used, the former is to “exclude those variants which don’t have SIFT and Polyphene deleterious score” and the second was “excluding those variants which don’t have minor allele frequency less than 5% (MAF ≤0.05). A total number 13 variants were selected which includes rs572205989, rs201814961, rs572178955, rs182452331, rs551202646, rs554705536, rs184929979 and rs573567598 (Figure 2).

We used the different free available databases of genomes to minimize and avoid FPR for the selection of True Positive Result (TPR) with their different criteria and algorithms. False Positive Result (FRP) prediction is a serious problem and we also found the difference in the number of variants (Figure 3). Venn diagram is used to show the intersection of results obtained from different databases (Figure 2).

Damaging nsSNPs Identifications

We checked and filter variants that were downloaded from different genome databases individually and found 152, 195, and 113 (UCSC genome browser, dbSNPs NCBI, and GnomAD database respectively) missense variants (Figure 4). These variants were only those that were predicted to be damaging and probably damaging (SIFT and PP2 respectively).

For further validation, different tools, namely SIFT, PolyPhene 2 (PP2Hvar and PP2Hdiv), Mutation Taster and, Mutation Tasting were used to predict phenotype-altering polymorphism (Table 1 & 2). Cross-validation is often not the best method to estimate actual performance (because leakage of information across multiple sources can lead to overly optimistic results) and time-stamped strategies are often preferable.

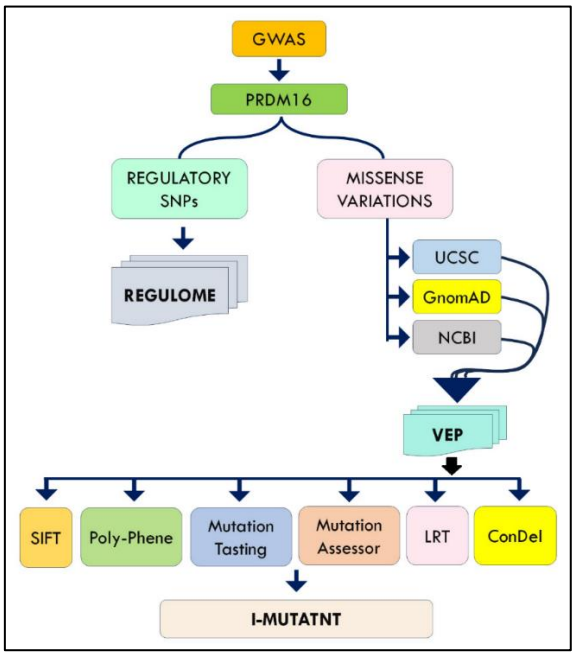

Figure 1: Schematic representation of computational tools for in silico analysis of PRDM16gene

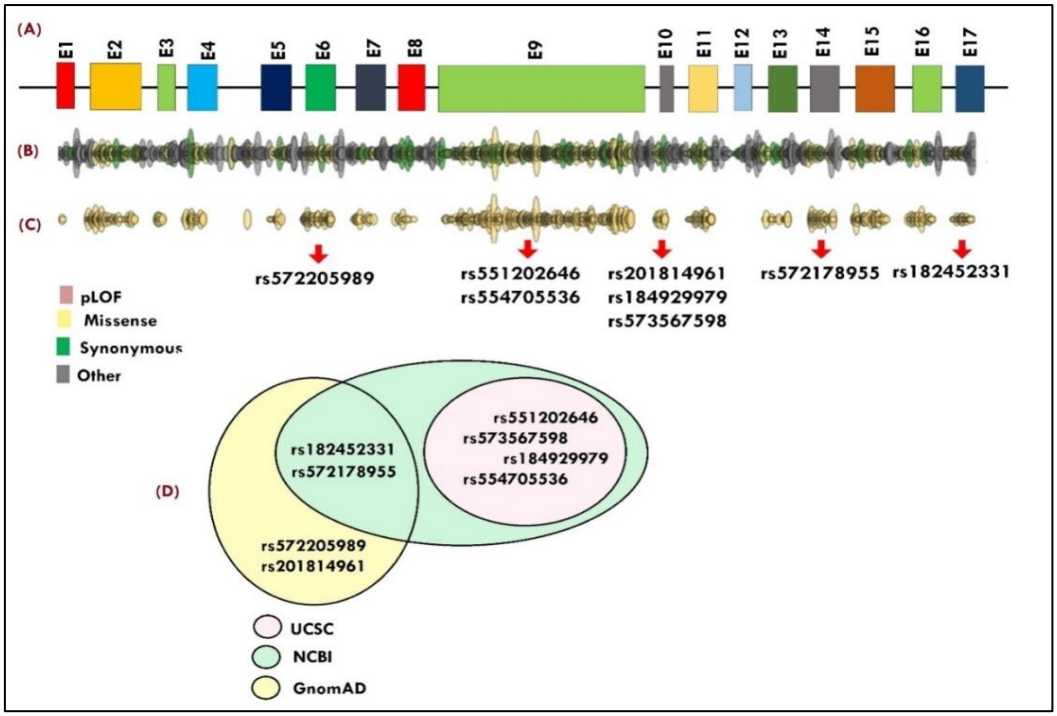

Figure 2: Graphical view of PRDM16 gene adapted from GnomAD database, (A):Picture representing the number of exons (17) of PRDM16 gene, (B): Different genomic variants which are colored with different color (peach color is Loss of function mutations, yellow color is Missense variants, light green depicts synonymous genetic variants and grey color are indel mutations), (C):Total filtered Missense variants from GnomAD database using VEP Ensembl, (D): Total filtered variants from GnomAD, NCBI variation viewer, and UCSC genome browser.

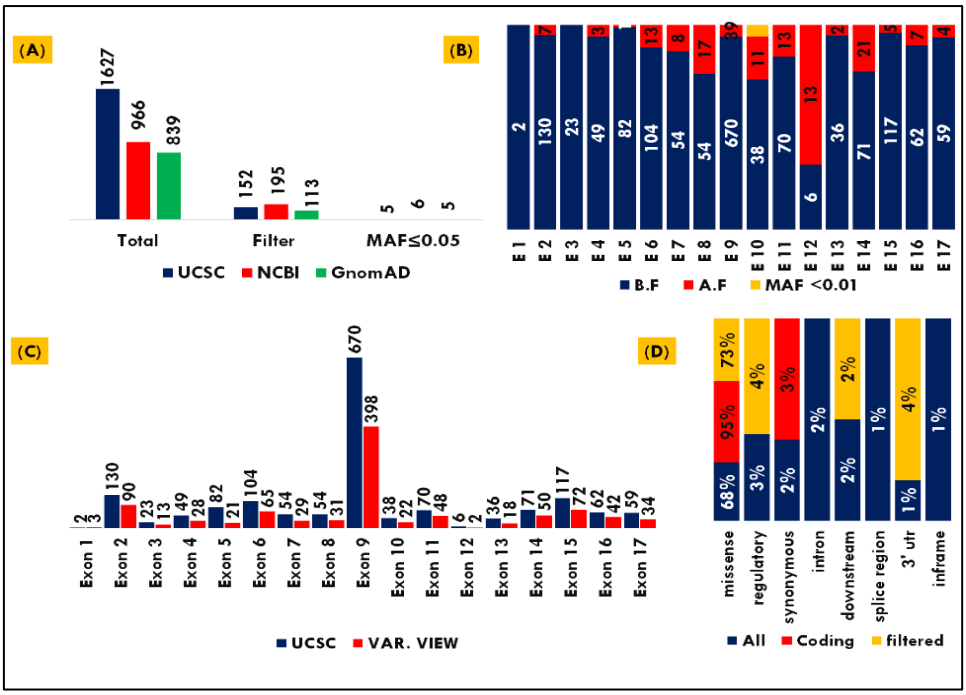

Figure 3: (A): Bar graph representing difference in total number of Missense variants downloaded from 3 different genome databases including UCSC genome browser (total variants 1627, after filtering found 152 missense variants and 5 missense variants having ≤0.05), NCBI (total variants 966, after filtering found 195 missense variants and 1 missense variants having ≤0.05, and GnomAD browser (total variants 839, after filtering found 113 missense variants and 5 missense variants having ≤0.05). (B): Bar graph represent the number of Missense mutations in PRDM16 gene with different exons downloaded from UCSC genome browser. The three different color coding represents the number of genetic variants before filtering (BF), after filtering Missense variants, and variants with MAF <0.01 using VEP Ensembl. (C): Difference in number of missense variants downloaded from UCSC genome browser and NCBI Variation viewer. (D): Bar graph represents different genetic variants from GnomAD browser with their percentage with inframe mutations (1%), 3' UTR (1%, with no variants in coding, and 4% after filtering), splicing region genetic variants with 1%, downstream genetic variants represent 2% and after filtering 2%, intronic variants 2%, synonymous variants 2%, coding region 3%, regulatory genetic variants with 3%, and 4% after filtering and missense variants 68%, with 95% in coding region and 73% after filtering (≤0.01).

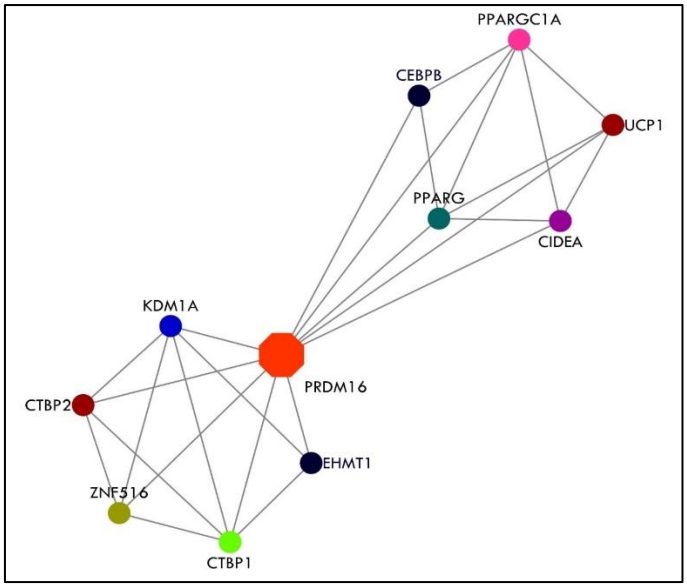

Figure 4: Protein-Protein Interaction of PRDM16 showed a significant interaction with different genes with an average node degree of 6.

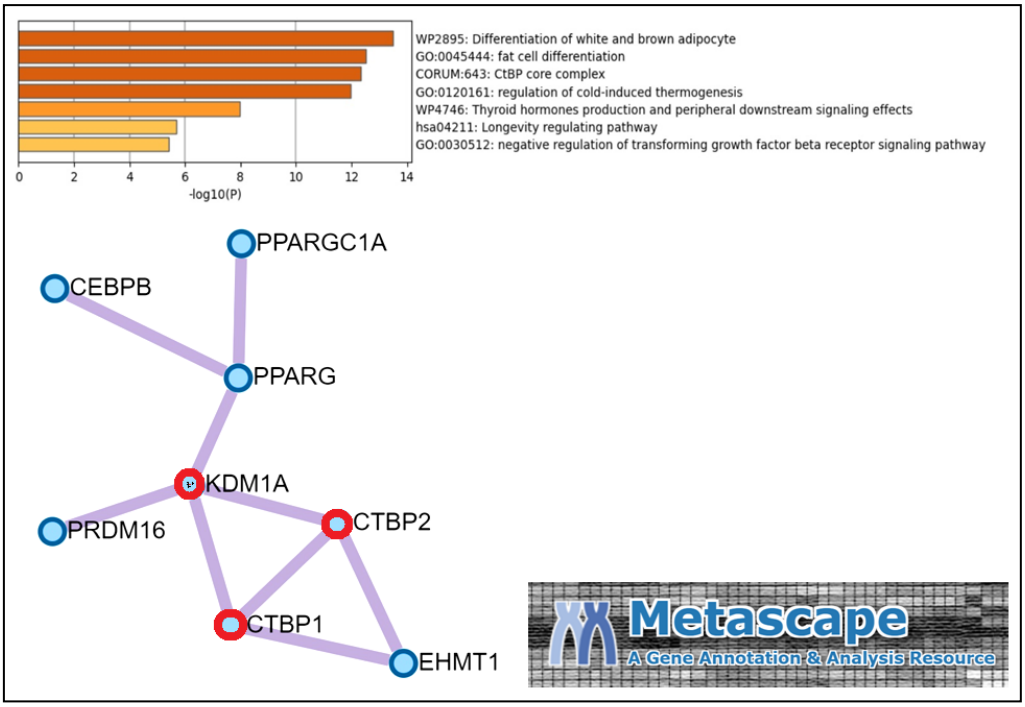

Figure 5: (Upper): MetaScape analysis showed increased expression in different process. (Below): Metascape Clustering showed two major cluster such as Cluster I=KDM1A, CTBP2, & CTBP1 and Cluster II= PRDM16, CEBPB, PPARGC1A, PPARG, & EHMT1.

Table 1: GnomAD browser & UCSC Filtered Variant (ENST00000270722.10)

Database	rs ID	Chr. loc. (Chr.1.)	Alt. allele	Exon	Codon changed	Amino acid
GnomAD	rs572205989	3402831	A	6	Tcc/Acc	S241T
	rs201814961	3414622	T	10	cGg/cTg	P889L
	rs572178955	3426162	T	14	tCg/tTg	S1074L
	rs182452331	3433777	G	17	gAg/gGg	E1266G
UCSC	rs551202646	3411478	G	9	gaC/gaG	D427E
	rs554705536	3411763	T	9	ttG/ttT	L522F
	rs184929979	3414574	A	10	cGg/cAg	R873Q
	rs573567598	3414628	T	10	cCg/cTg	P891L

Table 2: Variant prediction using different in-silico tools

Database	rs ID	VEST	SIFT	PP2HAR	PP2HDIV	Mutation Tasting	Mutation Assessor	LRT	ConDel
GnomAD	rs572205989	0.148	0.074=T	0.455=P	0.645=P	0.974=D	0.985=L	0.311448=U	0.554
	rs201814961	0.796	0.032=D	0.998=D	1.0=D	1=D	1.355=L	0.000102=D	0.88
	rs572178955	0.846	0.004=D	-	-	1=D	2.08=M	0.000124=D	0.85
	rs182452331	0.762	0.0=D	0.772=P	0.95=D	0.999=D	2.28=M	0.000175=U	0.732
UCSC	rs551202646	0.685	0.288=T	0.998=D	1.0=D	0.99=D	-1.465=N	0.000053=D	0.857
	rs554705536	0.442	0.044=D	0.988=D	1.0=D	0.99=D	2.295=M	0.001263=U	0.786
	rs184929979	0.697	0.001=D	0.992=D	1.0=D	0.99=D	1.975=M	0.000277=D	0.853
	rs573567598	0.695	0.0=D	0.965=D	1.0=D	1=D	1.65=L	0.000103=D	0.851

Table 3: Variant functional prediction values using I mutant

Database	rs ID	Strand	Amino acid	I-mutant	RI	SVM2 Prediction effect
GnomAD	rs572205989	+	S241T	-0.71 Kcal/mol	6	Decrease
	rs201814961	+	P889L	-1.19 Kcal/mol	8	Decrease
	rs572178955	+	S1074L	0.26 Kcal/mol	1	Increase
	rs182452331	+	E1247G	-	-	-
UCSC	rs551202646	+	D427E	-0.57 Kcal/mol	1	Decrease
	rs554705536	+	L522F	0.14 Kcal/mol	2	Increase
	rs184929979	+	R873Q	-0.19 Kcal/mol	3	Decrease
	rs573567598	+	P891L	-0.82 Kcal/mol	5	Decrease

Table 4: Regulatory Variants: Filtered Using Regulome database only 8 variants are found to be passed (ENST00000270722.10)

Chromosome location	dbSNP IDs	Location Intron	Rank	Score	1000 Genome Browser
chr1:3079166-3079167	rs79966817	1	2b	1.0	C=0.9187, T=0.08127
chr1:3079170-3079171	rs74943250	1	2b	0.76166	G=0.9187, T=0.08127
chr1:3226538-3226539	rs59127130	3	2b	0.97133	T=0.9319, A=0.06809
chr1:3229257-3229258	rs12409315	3	2b	0.63769	A=0.9171, G=0.08287
chr1:3229968-3229969	rs57323723	3	2a	1.0	G=0.9171, A=0.08287
chr1:3240724-3240725	rs118127661	3	2b	0.76026	T=0.9237, C=0.07628
chr1:3241290-3241291	rs57631188	3	2b	0.93883	T=0.9004, C=0.09964
chr1:3241293-3241294	rs12411233	3	2b	0.50489	C=0.9345, T=0.0655

4. DISCUSSION

Migraine, a polygenic, dysautonomic, complex neurological disorder with vascular dysfunction as consequences and is characterized by the "lower threshold of neuronal hyper-excitability" called "the migrainous brain" (Sudershan *et al*., 2023; Kogelman *et al*., 2020; Kumar *et al*., 2023). A meta-analysis of 22 GWA studies conducted by Gormley et al, including the data of a total of 59,674 affected subjects and 316,078 controls collected from six tertiary headache clinics and 27 population-based cohorts in which more than 35,000 new migraine cases that were not included in previously published GWA studies (Esserlind *et al*., 2013; Ligthart *et al*., 2011; Anttila *et al*., 2013; Freilinger *et al*., 2012; Anttila *et al*., 2010). These GWA studies identified 38 genomic loci harboring 44 independent susceptibility markers were identified in this GWAS for the prevalent forms of migraine (Gormley *et al*., 2016). PRDM16 rs2651899 were reported in the North Indian population associated with the migraine susceptibility (Ghosh *et al*., 2013) and act as a potential genetic marker for migraine susceptibility in MO (Migraine Without Aura) and female subgroup at both genotypic and allelic level in the North Indian population (Kaur *et al*., 2019).

In the present study, missense variants downloaded from 3 different genome databases including UCSC genome browser (total variants 1627, after filtering found 152 missense variants and 5 missense variants having ≤0.05), NCBI (total variants 966, after filtering found 195 missense variants and 1 missense variant having ≤0.05, and GnomAD browser (total variants 839, after filtering found 113 missense variants and 5 missense variants having ≤0.05) as shown in the (Figure 3A). Further detailed information about the data curation are given the Figure 3 such as missense SNP data of each exon of PRDM16 gene downloaded from UCSC showed number of missense mutation (Figure 3B), but after filtering (AF) significant decrease in number of missense mutation was observed (Figure 3B), and lastly using the filter of MAF <0.01 using VEP Insilco tool, only the 10th exon showed the missense mutation (Figure 3B). It was also observed that there were variation in the number of missense variants downloaded from the UCSC genome browser and NCBI's Variation viewer (Figure 3C). In addition to this, different genetic variants from GnomAD browser with their percentage with inframe mutations (1%), 3' UTR (1%, with no variants in coding, and 4% after filtering), splicing region genetic variants with 1%, downstream genetic variants represent 2% and after filtering 2%,

intronic variants 2%, synonymous variants 2%, coding region 3%, regulatory genetic variants with 3%, and 4% after filtering and missense variants 68%, with 95% in coding region and 73% after filtering (≤0.01) (Figure 3D).

Using the String DB (STRING: functional protein association networks (string-db.org) with the default setting such as Network type: Full String Network, minimum required interaction score: median confidence (0.400), the protein-protein interaction (PPI) was built, which was then edited by Cytoscape (Cytoscape: An Open Source Platform for Complex Network Analysis and Visualization). It was observed that there are 10 genes with which PRDM16 showed interaction with 33 number of edges, 6 average node degree, 0.807 average local clustering coefficient, and significant PPI enrichment p-value i.e., 2.07e-06 (Figure 4). Further the lists of interacting genes were further analyzed such as clustering using Metascape (Figure 5).

The current study may have some limitation including that these prioritized genetic variants have not been verified through experimental approach and the pathogenicity of the variants were also determined through a computational approach. So, the selected SNPs are required to be studied in the model organism and cell culture that may have potential importance in personalized medicines. Exploration of these prioritized genetic variants may provide novel remedial markers for various diseases and can be useful for association studies in clinical psychology and psychiatry, especially concerning the response to psychopharmacological and psychotherapeutic treatment.

False Positive Result (FRP) prediction is a serious problem and to minimize and avoid FPR for the selection of True Positive Result (TPR) we used different computational/online *in silico* tools with their different criteria and algorithms. Thus, only those SNPs that are predicted to be deleterious throughout these algorithms hold the potential for future evaluation and translation. Thus, combinations of varied computational tools come up with much more dimensions to predict the effect of genetic variants on the protein, which could be cost-effective.

5. CONCLUSION

In this study, we demonstrate a computational strategy for finding and predicting the functional impact of nsSNPs and amino acid substitution by using a combination of multiple in silico tools that provide more dimension to predict the effect of mutations on proteins. We found eight missense variants including rs572205989, rs201814961, rs572178955, rs182452331, rs551202646, rs554705536, rs184929979 and rs573567598that could play a role in the development of disease.

ABBREVIATION

SNPs: Single Nucleotide Polymorphisms
GWAS: Genome-Wide Association Studies
MA: Migraine with Aura
MWA: Migraine without Aura
PRDM16: PR domain containing 16
HMT: Histone Methyltransferases
UCSC: University of California Santa Cruz
VEP: Variant Effect Predictor
SIFT: Sorting Intolerant From Tolerant
PolyPhene2: Polymorphism Phenotyping v2
LRT: Likelihood Ratio Test
VEST: Variant Effect Scoring Tool
NCBI: National Center for Biotechnology Information
nsSNPs: Non Synonymous SNP
FPR: False Positive Result
TPR: True Positive Result

Contribution to the Paper

PK and AS conceptualized the study, AS wrote the manuscript, MB, SB, IS and RKP created all the pictures, graphs, and tables, and PK edited and finalized the manuscript.

CONFLICT OF INTEREST STATEMENT: The author declares that they have no conflict of interest.

FUNDING: No funding

Data Availability Statements

The datasets generated during and/or analyzed during the current study are available on different databases. References and web links are given in the References section of the current article.

LITERATURE CITED

- Adzhubei, I., Jordan, D. M., & Sunyaev, S. R. (2013). Predicting functional effect of human missense mutations using PolyPhen-2. *Current protocols in human genetics*, *76*(1), 7-20.
- Andreou, A. P., & Edvinsson, L. (2019). Mechanisms of migraine as a chronic evolutive condition. *The journal of headache and pain*, *20*(1), 117. https://doi.org/10.1186/s10194-019-1066-0
- Anttila, V., Stefansson, H., Kallela, M., Todt, U., Terwindt, G. M., Calafato, M. S., Nyholt, D. R., Dimas, A. S., Freilinger, T., Müller-Myhsok, B., & Artto, V. (2010). Genome-wide association study of migraine implicates a common susceptibility variant on 8q22.1. *Nature genetics*, *42*(10), 869.
- Anttila, V., Winsvold, B. S., Gormley, P., Kurth, T., Bettella, F., McMahon, G., Kallela, M., Malik, R., De Vries, B., Terwindt, G., & Medland, S. E. (2013). Genome-wide meta-analysis identifies new susceptibility loci for migraine. *Nature genetics*, *45*(8), 912-917.
- Bhatti, P., Church, D. M., Rutter, J. L., Struewing, J. P., & Sigurdson, A. J. (2006). Candidate single nucleotide polymorphism selection using publicly

available tools: a guide for epidemiologists. *American journal of epidemiology*, *164*(8), 794-804.

- Bolay, H., Reuter, U., Dunn, A. K., Huang, Z., Boas, D. A., & Moskowitz, M. A. (2002). Intrinsic brain activity triggers trigeminal meningeal afferents in a migraine model. *Nature medicine*, *8*(2), 136-142. https://doi.org/10.1038/nm0202-136
- Boyle, A. P., Hong, E. L., Hariharan, M., Cheng, Y., Schaub, M. A., Kasowski, M., ... & Snyder, M. (2012). Annotation of functional variation in personal genomes using RegulomeDB. *Genome research*, *22*(9), 1790-1797.
- Boyle, E. A., Li, Y. I., & Pritchard, J. K. (2017). An expanded view of complex traits: from polygenic to omnigenic. *Cell*, *169*(7), 1177-1186.
- Capriotti, E., Fariselli, P., & Casadio, R. (2005). I-Mutant2. 0: predicting stability changes upon mutation from the protein sequence or structure. *Nucleic acids research*, *33*(suppl_2), W306-W310.
- Carter, H., Douville, C., Stenson, P. D., Cooper, D. N., & Karchin, R. (2013). Identifying Mendelian disease genes with the variant effect scoring tool. *BMC genomics*, *14*(3), 1-16.
- Chun, S., & Fay, J. C. (2009). Identification of deleterious mutations within three human genomes. *Genome research*, *19*(9), 1553-1561.
- Cooper, D. N. (1992). Regulatory mutations and human genetic disease. *Annals of medicine*, *24*(6), 427-437.
- Esserlind, A. L., Christensen, A. F., Le, H., Kirchmann, M., Hauge, A. W., Toyserkani, N. M., ... & Olesen, J. (2013). Replication and meta-analysis of common variants identifies a genome-wide significant locus in migraine. *European journal of neurology*, *20*(5), 765-772.
- Freilinger, T., Anttila, V., De Vries, B., Malik, R., Kallela, M., Terwindt, G. M., ... & International Headache Genetics Consortium. (2012). Genome-wide association analysis identifies susceptibility loci for migraine without aura. *Nature genetics*, *44*(7), 777-782.
- Ghosh, J., Pradhan, S., & Mittal, B. (2013). Genome-Wide-Associated Variants in Migraine Susceptibility: A Replication Study From North I ndia. *Headache: The journal of head and face pain*, *53*(10), 1583-1594.
- Gnad, F., Baucom, A., Mukhyala, K., Manning, G., & Zhang, Z. (2013). Assessment of computational methods for predicting the effects of missense mutations in human cancers. *BMC genomics*, *14*(3), 1-13.
- Gormley, P., Anttila, V., Winsvold, B. S., Palta, P., Esko, T., Pers, T. H., ... & Palotie, A. (2016). Meta-analysis of 375,000 individuals identifies 38 susceptibility loci for migraine. *Nature genetics*, *48*(8), 856-866.
- Kaur, S., Ali, A., Ahmad, U., Pandey, A. K., & Singh, B. (2019). rs2651899 variant is associated with risk for migraine without aura from North Indian population. *Molecular biology reports*, *46*(1), 1247-1255.
- Kogelman, L. J., Falkenberg, K., Buil, A., Erola, P., Courraud, J., Laursen, S. S., Michoel, T., Olesen, J., & Hansen, T. F. (2020). Changes in the gene expression profile during spontaneous migraine attacks. *bioRxiv*.
- Kumar, S., Raina, J. K., Sudershan, A., Mahajan, K., Jasrotia, R., Maharana, C., ... & Kumar, P. (2023). An Association Study of ESR 1–Xba I and Pvu II Gene Polymorphism in Migraine Susceptibility in the Jammu Region. *European Neurology*, *86*(1), 55-62. https://doi.org/10.1159/000527271
- Leao, A. A. (1944). Spreading depression of activity in the cerebral cortex. *Journal of neurophysiology*, *7*(6), 359-390. https://doi.org/10.1152/jn.1944.7.6.359
- Ligthart, L., De Vries, B., Smith, A. V., Ikram, M. A., Amin, N., Hottenga, J. J., Koelewijn, S. C., Kattenberg, V. M., De Moor, M. H., Janssens, A. C. J., & Aulchenko, Y. S. (2011). Meta-analysis of genome-wide association for migraine in six population-based European cohorts. *European Journal of Human Genetics*, *19*(8), 901-907.
- Manolio, T. A. (2010). Genomewide association studies and assessment of the risk of disease. *New England journal of medicine*, *363*(2), 166-176.
- Ng, P. C., & Henikoff, S. (2003). SIFT: Predicting amino acid changes that affect protein function. *Nucleic acids research*, *31*(13), 3812-3814.
- Patnala, R., Clements, J., & Batra, J. (2013). Candidate gene association studies: a comprehensive guide to useful in silico tools. *BMC genetics*, *14*, 1-11.
- Reva, B., Antipin, Y., & Sander, C. (2011). Predicting the functional impact of protein mutations: application to cancer genomics. *Nucleic acids research*, *39*(17), e118-e118.
- Schwarz, J. M., Cooper, D. N., Schuelke, M., & Seelow, D. (2014). MutationTaster2: mutation prediction for the deep-sequencing age. *Nature methods*, *11*(4), 361-362.
- Sladek, R., Rocheleau, G., Rung, J., Dina, C., Shen, L., Serre, D., ... & Froguel, P. (2007). A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature*, *445*(7130), 881-885.
- Sudershan, A., Bharti, S., Sheikh, J. H., Wani, S. A., Kumar, H., & Kumar, P. Epigenetics in migraine: A review. https://doi.org/10.18231/j.ijn.2023.026
- Sudershan, A., Mahajan, K., Singh, K., Dhar, M. K., & Kumar, P. (2022). The complexities of migraine: a debate among migraine researchers: a review. *Clinical neurology and neurosurgery*, *214*, 107136. https://doi.org/10.1016/j.clineuro.2022.107136

- Sudershan, A., Pushap, A. C., Younis, M., Sudershan, S., Bhagat, S., Kumar, H., Panjalyia, R. K., & Kumar, P. (2023) Neuroepidemiology study of headache in the region of Jammu of north Indian population: *A cross-sectional study*. *Front. Neurol*, *13*, 1030940. doi: 10.3389/fneur.2022.1030940
- Sudershan, A., Sudershan, S., Younis, M., Bhagat, M., Pushap, A. C., Kumar, H., & Kumar, P. (2023). Enlightening the association between TNF-α-308 G> A and migraine: a meta-analysis with meta-regression and trial sequential analysis. *BMC neurology*, *23*(1), 1-14. https://doi.org/10.1186/s12883-023-03174-x
- Sudershan, A., Younis, M., Sudershan, S., & Kumar, P. (2023). Migraine as an inflammatory disorder with microglial activation as a prime candidate. *Neurological Research*, *45*(3), 200-215. https://doi.org/10.1080/01616412.2022.2129774
- Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., ... & Mering, C. V. (2019). STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic acids research*, *47*(D1), D607-D613.
- Tam, V., Patel, N., Turcotte, M., Bossé, Y., Paré, G., & Meyre, D. (2019). Benefits and limitations of genome-wide association studies. *Nature Reviews Genetics*, *20*(8), 467-484.
- Timpson, N. J., Greenwood, C. M., Soranzo, N., Lawson, D. J., & Richards, J. B. (2018). Genetic architecture: the shape of the genetic contribution to human traits and disease. *Nature Reviews Genetics*, *19*(2), 110-124.
- Tranchevent, L. C., Ardeshirdavani, A., ElShal, S., Alcaide, D., Aerts, J., Auboeuf, D., & Moreau, Y. (2016). Candidate gene prioritization with Endeavour. *Nucleic acids research*, *44*(W1), W117-W121.
- Warde-Farley, D., Donaldson, S. L., Comes, O., Zuberi, K., Badrawi, R., Chao, P., ... & Morris, Q. (2010). The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic acids research*, *38*(suppl_2), W214-W220.
- Wray, N. R., Wijmenga, C., Sullivan, P. F., Yang, J., & Visscher, P. M. (2018). Common disease is more complex than implied by the core gene omnigenic model. *Cell*, *173*(7), 1573-1580.