Structural, functional, and evolutionary level mutational analysis of TREM2 gene to study its role in Alzheimer's disease

Enrollment No – 207803

Name – Ajay Kumar

Supervisor – Dr. Tiratha Raj Singh

May-2022

Submitted in partial fulfillment of the Degree of

Master of Science

in

Biotechnology

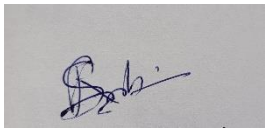
DEPARTMENT OF BIOTECHNOLOGY AND BIOINFORMATICS

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY

WAKNAGHAT, SOLAN, -173234, H.P., INDIA

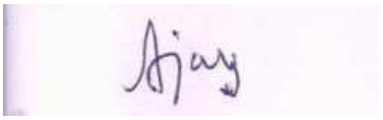
CERTIFICATE OF ORIGINALITY

This is to certify that the thesis titled **Structural, functional, and evolutionary level mutational analysis of TREM2 gene to study its role in Alzheimer's disease** is an original work of the student and is being submitted in partial fulfilment for the award of the Degree of **Master of Science (Biotechnology)**. This dissertation thesis has not been submitted earlier either to this University or to any other University/Institution for the fulfilment of the requirement of any course of study.



Signature of Supervisor

Date: 30/05/2022



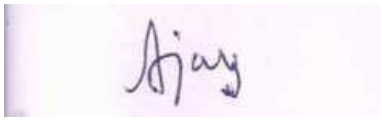
Signature of Student

Date: 30/05/2022

UNDERTAKING FROM THE CANDIDATE

I**, Ajay Kumar,** present the project entitled **"Structural, functional, and evolutionary level mutational analysis of TREM2 gene to study its role in Alzheimer's disease".** Though care has been taken while writing this report, there may be still some errors (typographical or otherwise) which are inadvertent on my part.

Place: **Jaypee University of Information Technology, Waknaghat, Distt. Solan, H.P.**



Signature of Candidate

Ajay Kumar

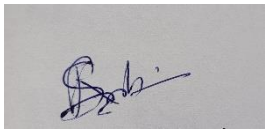
CERTIFICATE OF UNIVERSITY FACULTY GUIDE

Enrolment No: 207803

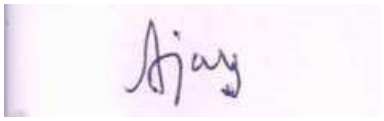
This is to certify that **Mr. Ajay Kumar** of M.Sc. (Biotechnology) has completed this Research/Dissertation Project under my supervision in partial fulfilment for the award of the Master of Science Degree in (Biotechnology) from **Jaypee University of Information Technology, Waknaghat, Distt. Solan, Himachal Pradesh**.

Course Name: **DISSERTATION**

Course Code: **20MS9BT491**



SIGN OF FACULTY GUIDE



SIGN OF STUDENT

PLACE: **Jaypee University of Information Technology, Waknaghat, Distt. Solan, H.P.**

DATE: 30/05/2022

ACKNOWLEDGEMENT

I would like to express my profound gratitude to my guide **Dr. Tiratha Raj Singh Sir** for his guidance, support, and constant encouragement throughout the course of this project work. He has been more than just my project guide; at times a mentor to rescue me out of my doubts, at times a philosopher to hear my miseries, and still at times a friend with whom I can share myself. He has always pushed me to work hard and still gave me enough space to cultivate and implement my ways of dealing with the problems. Moreover, he taught me life values that I will never forget.

Furthermore, I would like to acknowledge Vice-Chancellor Prof. (Dr.) Rajendra Kumar Sharma Sir, Prof. (Dr.) Ashok Kumar Gupta Sir, Dean of academics & research for providing me with an opportunity to be a part of the institute and to complete my Master's Degree.

I also want to mention the HOD of Biotechnology and Bioinformatics Prof. (Dr.) Sudhir Syal Sir has been a source of immense motivation and inspiration both for my academic and personal life. He was never, and I know will never be, more than just a phone call away. With his acquaintance of Omni-domain knowledge and skills, he has helped me in almost every aspect I have asked him for.

In addition, I would like to thank all the faculty members of the BT&BI Department of JUIT, who have mended a novice in the field of Bio-Technology into a well-equipped bio-technologist. Adding to the list, I would like to thank all the lab engineers and Ms. Somlata Sharma ma'am for providing me with a workspace.

Also, I would like to thank Mr. Rohit Shukla and Mr. Arvind Kumar who have always been available to help and provide me with helpful insights on my work. A few other Ph.D. scholars have been a great help in my work.

I would also like to thank Ms. Rhea Singh for her constant support, for being there for me, motivating, caring, and always cheering me up whenever life knocks me down.

I would also like to appreciate the part that my classmates have played in shaping this project work. They have been my constant support and cheered me up at hard times which has helped me a lot. Thanks a lot!

A big thanks to all open-source tool providers because of which I was able to complete my project.

I would like to thank the almighty God for his grace throughout my life. Last but not least I would like to thank my Mother and Father who have always supported me through thick and thin and have been a constant source of encouragement and support; also, my younger brother who has never given up on me and always motivated me.

I Thanks to **JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY**

AJAY KUMAR

M.Sc. Biotechnology

JUIT, Solan

LIST OF TABLES

LIST OF FIGURES

INDEX

TOPICS	PAGE NUMBER

ABSTRACT

The TREM 2 gene present in the myeloid cell increases the risk of Alzheimer's disease (AD) in an individual. TREM 2 cells can be associated with the cause of inflammation and immune disorder which is a major cause of AD. Single-nucleotide polymorphisms (SNPs) associated with complex diseases can originate, remove, or change protein coding sites. TREM2 stands for Triggering Receptor Expressed on Myeloid Cells 2, a transmembrane receptor that regulates the activity and survival of microglia. Some of the rare variants are associated with an increased risk of AD related to TREM2. A systematic in-silico investigation of disease-associated nsSNPs in human TREM2 was performed in this study. We have retrieved the nsSNPs from the NCBI's dbSNP and then several sequences (PROVEAN, PolyPhen-2, Meta-SNP, PMut, SNAP2, PhD-SNP, Suspect) and structure-based tools (DynaMut, PON-P2, SDM, Align GVGD, SNP&GO 3D, MUPro, DUET, iStable, I-Mutant) were used to screen out the most affecting SNPs. We have finally predicted six extremely deleterious nsSNPs (Y38C, Y38H, W50S, C51G, D104G, andV124D). Most nsSNPs are found in highly conserved residues according to the conservation analyses using ConSurf. Additional computational studies also support this view and results for all these analyses have been presented for this study. The Molecular dynamic simulation analysis revealed that all the mutations are altering the proper geometry of the TREM2 protein structure. This computer study is expected to aid in the prioritization of genes and SNPs, explaining the genetic pathways behind neuroinflammation and neurodegenerative diseases including AD.

Keywords - Single nucleotide polymorphisms (SNPs), TREM2, Neuroinflammation, Neurodegenerative Disease, Alzheimer's Disease, Molecular dynamic simulations (MDS)

CHAPTER:1. INTRODUCTION

1.1. NEUROINFLAMMATION

Neuroinflammation is the inflammatory response incurred in the brain & its spinal extensions. Immuno-entities like cytokines, chemokines, reactive oxygen species (ROS), and secondary messengers have a significant role in neuroinflammation. Major neuro-companion cells, endothelial cells, and peripherally derived immune cells are responsible for immune mediation. Immunological, physiological, metabolic, and psychological ramifications are well attributed to neuroinflammation. Microglia play a major role in neuroinflammation. This is because these innate immune cells are responsible for the CNS's primary immunological surveillance and macrophage-like functions, such as the production of cytokines and chemokines. The majority of the CNS's innate immunological capabilities are because of microglial cells. The white and grey matter of the brain and spinal cord do contain major content of resident CNS cells. 10% of the CNS population is comprised of microglia[1].

Trauma, infection, tau oligomers, oxidative agents, Aβ, and redox iron all appear to be neuroinflammation-causing signals. Aberrant release of proinflammatory cytokines is caused by neuroinflammation, which activates signalling mechanism that causes brain tau hyperphosphorylation in residues that are not changed under normal physiological settings. Indeed, evidence exists that AD pathogenesis involves significant interactions with immunological cells in the brain, such as glial cells, and invading immune cells from the periphery, which may help to modify the neuroinflammation and neurodegenerative process in AD brains. The Neuroimmunomodulation hypothesis plays a major role in this context, concentrating on the relationship between damage in neurons and the inflammation in the brain. Activation of astrocyte and microglial cells push the condition to the level resulting in the overproduction of chemicals responsible for underlying neuroinflammation[2].

The role of the innate immune system in disease development has been demonstrated to have a negative bidirectional relationship with tau pathology. The tau protein is a member of the MAPs family and is primarily expressed by neurons with a preference for axonal location. Tau has been found to increase tubulin polymerization and reduce the transition rate between growth and contraction phases in microtubules in vitro, resulting in a stable but dynamic state[3], [4]. Tau is mostly present in axons, although a tiny quantity is also found in dendrites. Tau's post-synaptic function is unknown; however, it

may play a role in synaptic plasticity[5]. In contrast, a nuclear role of tau has been revealed, which might involve in maintaining transcriptional activity and preserving DNA/RNA integrity in both normal and stressful situations.[6]

The structure of tau represents a graded phosphorylation mechanism in which distinct locations affect the protein's orientation, allowing secondary kinases to function more effectively. Different parts of the brain are phosphorylated earlier than others in Alzheimer's disease, causing in the creation of novel surface proteins. [7], [8]. It was also shown that microglial cell production of tau enhances tau activation[9]. The exact pathway leading to tau phosphorylation is still not known, Tau gets separated from microtubules from subsequent structural modifications and creates more soluble free tau. Dynamic and gradual self-assembly supported by tau hyperphosphorylation in oligomeric aggregates and insoluble substances like as PHFs throughout the illness with neurotoxicity of varying degree before production of NFTs[10]. Injury, infection, toxin exposure, neurodegenerative illness, and aging all stimulate the brain's innate immune system[11].

In a recent GWAS, about 20 gene variants were identified that can confer a higher risk of LOAD. These genes include (CR1, ABCA7 CLU, HLA DRB5, HLA-DRB1, EPHA1, CD33, MS4A), endocytosis (PICALM, BIN1, EPHA1 and CD2AP, SORL1), and lipid biology (ABCA7, CLU, ABCA7, APOE). AD is defined as a decline in memory power and loss of function. To analyse it neuropathologically, it can be seen as a classic case of absorption of Aβ protein and the dilation of the ventricular and the webbed intraneuronal series of tau (t) proteins. These proteins activate the microglia and astrocytes which then mobilize and gather around, and release the inflammation-causing protein in the brain[12].

CHAPTER:2. REVIEW OF LITERATURE

2.1. TREM2

If we look closely at the studies that talk about the formation of Alzheimer's, one would understand the complexity of the disorder. There is a mutation in the genes namely the presenilin-2, PSEN1, and the amyloid precursor protein (APP) which is highly penetrant and mutation of the common genetic codes that decrease the risk of the disease and the moderate frequency of E4 allele of the apolipoprotein E (APOE) puts one insignificant risk of the disease. The presence of the gene variant with frequency lesser than and or same as APOE ε4 and their effect on patients still is a debatable topic[13] Analysis of the whole genome leads to the decoding of a rare variant that is present in the myeloid cell 2 known as TREM 2 and can be considered to be a triggering factor in the causing of Alzheimer's. It is responsible for moderating and controlling inflammation[14].

2.1.1 Structure of TREM 2

The gene type belongs to the TREM family and is located in the human chromosome spreading to a length of 4676 bp and is expressed widely in the myeloid cells. These gene types can also be expressed in the monocytes, dendritic cells, and macrophages[15]. The feature of the family of this genotype is that they have immunoglobulin domains their surface is covered with glycoprotein and they have a specified trans-membrane with a short cytoplasmic tail. The membrane is also known for activation of the adaptor protein like the DNAX- activation protein also known as DAP12 and the DAP10. The function of DAP 12 is that helps in the transmission of the cell through the recruitment of tyrosine protein and DAP10 promotes the functioning of the phosphatidylinositol. When functioning together with DAP 12 Mu and DAP 10, TREM 2 results in the mobilization of Ca^{2+}, the rearrangement of the cytoskeleton, targeting of the rapamycin, and the mitogen-activated protein kinase and is also associated with the activation of metabolism[16]. Concerning the functioning of the brain, TREM 2 is responsible for controlling the inflammatory responses and the process of phagocytosis of the debris. It was earlier found in a study that TREM 2 plays a significant role concerning the regulation of the myeloid cell, but the actual effect of myeloid cells in Alzheimer's disease is still not found. If found in the primary microglia it can reduce the number of cells. It has been studied that neuroinflammation is directly affected by an increasing number of myeloid cells as such[17].

2.1.2. TREM Signalling

In opposition to the methods used to discover the TREM-1 or TREM-2 ligands, the TLT-2 receptor, the only completely described TREM ligand discovered too far, was discovered using a novel process. Hashiguchi et al. observed considerable similarities among the CD28 family, recognized receptors for other B7 compounds, and TLT-2 when trying to uncover receptors for the B7 member of the family B7-H3. This led them to examine B7-H3 affinity to TLT-2[15]. Bioinformatics techniques similar to those used by Hashiguchi and coworkers may be helpful for future efforts to find TREM ligands.

2.1.3. TREM 2 and Neuroprotection

In vitro study showed that, due to oxidative stress, Aβ can negatively interfere with the synaptic transmission and leads to cell death. Likewise, microglia in the brain have already been found to consume Aβ. As a result, phagocytosis by microglia may have a neuroprotective effect. (18) TREM2 deficiency, on the other hand, drastically reduces microglia's capacity to ingest amyloid plaques. As per the study conducted on the animal AD model, those that were TREM2 - deficient showed a huge decline in the concentration of microglia around amyloid[18].

A loss of TREM2 or DAP12 in AD animal models causes more distributed neuritic plaque and more synapses, resulting in a topology that mimics a sea urchin. The degree of harm to the nervous system depends on the count of the synapses that come through outside, if the number is greater than the surface contact with the nerve structure is greater and so is the injury. The amyloid plaques are separated from the peripheral nerve tissues by the microglia which acts as a barrier between the two and also plays a protective role in the brain[19].

2.1.4. TREM2 and modulation of inflammatory responses

Whereas most studies found the anti-inflammatory effect of TREM2 and complicated linkage between TREM2 and other inflammatory reactions. The intensity and duration of the stimuli variably depend on the cell and situation.

In systemic inflammation, TREM2 appears to perform variable roles. In several *in vitro* and *in vivo* investigationsTREM2 has been found to have an anti-inflammatory function. The decrease in the number of TREM2 also has an effect on the increase of the proinflammatory mediators like the interleukin-1 (IL1) necrosis factor (TNF), and interleukin-6 in cell lines (IL6) [20].TREM2 reduction

boosted the release of inflammatory cytokines in the lab test conducted on the mouse P8 (SAMP8) animal model. Additionally, in AD mouse models, overexpression of TREM2 lowered the amounts of proinflammatory genes. One can deduce from these findings that TREM2 can reduce inflammatory responses in specific situations. Furthermore, due to the paucity of these variations, the possibility of them having an influence on risk prediction at the population level is minimal, although polygenic risk forecasts for AD could still be having a positive influence by adding TREM2 mutations. Except for p.Arg47His and p.Arg62His, none of these variations cause disease[21].

2.1.5. TERMS 2 and association with Alzheimer's disease

There have been 46 genetic variations found in TREM2 that have been examined for AD so far. One of the many genetic variations, p. Arg47His (rs75932628), is extremely rare. — And has been linked to higher risk and increased the chances by up to two to three times of having Alzheimer's illness[22]. This has been observed in studies conducted on several European and North American peoples. There is an important connection between rs75932628 and a proxy. TREM2's influence appears to be population-based.

In kinship datasets, p.Arg47His and the risk of Alzheimer's disease showed moderate association[23]. In a research of 130 households with LOAD, p.Arg47His was said to be found in four (3%) families (10 times compared with the general population), and in a big family of 21 people with delayed AD (16 of whom had accessible genetic information), (75%) persons with AD borne the p.Arg47His allele, noted with p.Arg disease has shown vertical transmission[24].

Many TREM2 variations have been examined for their link well with the probability of AD (Alzforum provides an updated participatory presentation of TREM2 variants), with only p.Arg62His displaying genomic affiliation levels of importance, irrespective of the p.Arg47His correlation. In a study conducted by the Chinese two more polymorphisms were discovered namely the p.Ala130Val22 and p.Ala192Thr30[25].

These were majorly seen in individuals with LOAD. The p.His157Tyr and p.Ser183Cys variations are more common among people suffering from the mentioned disease. The p.His157Tyr genotype is strongly related to an elevated risk of LOAD, opposite to what had been discovered in adults of the European population. Further studies have provided proper evidence on the identification of amino acids involved in AD [26].

Additional mutations in the TREM gene cluster are linked to the risk of Alzheimer's disease or modification of the TREM gene expression. Other than the TREM2 variations, another layer of the allele for AD was discovered, which was positioned from TREML2 about 55kb downstream and upstream for about 24kb (rs9381040), held in nucleotide polymorphism[27].

A TREML2 missense variant, p.Ser144Gly (rs3747742), has also been linked to a reduced risk of AD. The link between these two polymorphisms and a lower risk of AD cannot be shown to exist independently with p.Arg47His genotype. Furthermore, rs9357347, hypersensitivity of DNase- area between TREM2 and TREML2 genotypic forms, has been associated with a lower risk of AD. This variation linkage conflicts with rs9381040 and subsequent investigation revealed that the rs9357347 variation serves as both a regulator and a functional executor. When these genotypes were associated with TREM1 and TREM2 and mRNA, here number was said to be seen to be increased[28].

Multiple sclerosis (MS) is by far the most prevalent central nervous system illness in adults, resulting in neurological damage and dysfunction. The potential for long-term axonal degeneration is caused by chronic inflammation in the CNS (Goldenberg, 2012). Patients can be categorized.80-90 percent of all cases begin with and are grouped according to the course of the disease[17]. The course is relapsing-remitting. The incidence of relapses decreased over time, and patients were able to return to work. The condition is in the second time gets, which implies it is becoming worse. In 10 minutes, MS begins in 20% of individuals as a primary progressive disease with no relapses or remissions. When compared to other forms of MS, this one is more resistant to treatment[29].

The exact etiology of the autoimmune condition is unknown; however, it is thought that several genetic vulnerabilities-genetic effects such as environmental parameters, infectious agents, and metabolism can cause it (Ramagopalan & Sadovnick, 2011)[30]. The discovery of a genetic disposition to MS has sparked a flurry of research into disease locus and alleles. About 100 genes have been discovered in genome-wide research, all of which appear to have a role in MS. Demyelination in parts of the brain and spine is the key pathological hallmark of MS development. Infiltrating T- and B-cells, monocytes, and activated microglia have been seen in locations with myelin degeneration, according to studies[31].

In MS, engaged microglia are observed in large numbers in white matter lesions that are demyelinating. The exact cause of microglia activity is unknown, while numerous theories have been offered. Microglia can be activated by pro-inflammatory cytokines generated by T and B cells, as well as by the phagocytosis of myelin and neurodegeneration debris[32]. In addition, pathogen stimulation of the TLR

pathway is thought to activate microglial, which is consistent with other studies that propose infectious agents have a role in MS (Sriram, 2011). Microglia stimulation could be detrimental or protective in MS. Several investigations have shown that several of the mechanisms that promote microglia are harmful in demyelinated lesions[33].

The recruitment and the reactivation of the T-cell can be done in various ways some of which are secretion, proteases, and the creation of inflammation-causing cytokines and reactive oxygen species, as well as toxicity to neurons and oligodendrocyte precursor cells, all of which would play a role in the worsening of MS progression (Rawji & Yong, 2013) [34]. Microglia may, however, play a helpful role in MS, according to new data. The three mechanisms that underpin these positive functions can be split. The first important function lies in the production of cytokines, chemokine, and growth factors that fasten the process of repair and the promotion of demyelination. The second function is that microglia can phagocytosis inhibiting debris and removal of the apoptotic cells, and third, stem cell populations can be recruited and neurogenesis can be induced [35].

Using their manufactured monoclonal antibody targeting mouse TREM-2 receptor, Piccio and colleagues investigated TREM-2 protein levels on microglia in EAE animals in 2007. TREM2 production was barely discernible in the spinal cords during the initial inflammatory stages of EAE, but it was elevated and remains higher in the long-term (later) phase[36]. A similar expression pattern also was observed in the DAP12 transcript and brain transcript analyses. TREM-2 expression was found to be instantly unregulated 18 days after vaccination, which corresponds to the peak of inflammation in the disorder[37]. They investigated TREM-2 protein expression using immunofluorescence and discovered high and diffuse utterance 20 days after microglia/macrophages and dendritic cells immunization on the microglia/macrophages and dendritic cells immunization[38].

CHAPTER:3. METHODOLOGY

3.1. DATA COLLECTION

The information about Human TREM2 nsSNPs (Gene Id **54209**) was obtained from the NCBI dbSNP database (http://www.ncbi.nlm.nih.gov/SNP/) (Date of access: October 14, 2021). 242 missense nsSNPs were preferred for screening. The UniProt database(http://www.uniprot.org) was Utilized for extraction of the human TREM2 sequence (UniProt ID: **Q9NZC2**). The Protein Data Bank was used to obtain the TREM2 protein structure (PDB ID: 5ELI). TREM2 gene has a total length of 4676 base pairs which is situated on chromosome number 6, from 41,126,246 bp to 41,130,922 bp.

3.2. DELETERIOUS nsSNPsIDENTIFICATION USING SEQUENCE BASED PREDICTION TOOLS

The functional implications of the nsSNPs related to protein were predicted utilising 9 sequence based tools. Protein Variation Effect Analyzer (PROVEAN) and Polymorphism Phenotyping v2 (PolyPhen -2**)** were used to quantify the impact of amino acid (AA) substitution related to the structural and functional level of proteins. Other consensus approaches like Meta-SNP and PredictSNP were used as well. The disease association and pathogenicity related to the nsSNPs were also predicted using SuSPect, PMut, SNAP2, PhD-SNP, and PON-P2.

3.2.1. PROVEAN

http://provean.jcvi.org/index.php

Is a tool that generally utilize an alignment based score to estimate if a protein's biological activity is affected by amino acid alteration[39]. The sequence similarity of a query sequence to a protein sequence homolog is compared with score without and with a change in amino acid within the query sequence. Protein variation is expected to have a "deleterious" impact if the Protein Variation Effect Analyzer score is $\leq$-2.5, and a "neutral" effect if the score is >-2.5.

3.2.2. PolyPhen -2

http://genetics.bwh.harvard.edu/pph2/

Based on sequence-based characterisation, PolyPhen -2 tool was used to assess the functional consequences of amino acid substitution on structure and function of Proteins. The prediction result was obtained in the form of a likelihood score that categorizes the variants seems predictably damaging, potentially damaging, and benign'. As a query, WT residue along with the position, and different amino acids of screened nsSNPs were submitted, along with the UniProt protein accession ID**Q9NZ**[40].

3.2.3. SNAP2

https://rostlab.org/services/snap2web/

SNAP2 works on the principle of the neural network categorization approach. It predicts how nsSNPs will affect the function of a protein. FASTA protein sequence and lists of mutants were submitted as input queries, with scores for each substitution that could be converted into neutral or non-neutral impact binary predictions[41].

3.2.4. PhD-SNP

http://snps.biofold.org/phd-snp/

This tool was utilized for determine if the single point mutation of protein is disease-related or neutral. It is based on support vector machines (SVM), which can check all information from current databases regarding variants [42]. The protein sequences were submitted as input to the PhD-SNP for examination, along with the residue changes.

3.2.5. PMut

(http://mmb2.pcb.ub.es:8080/PMut)

To predict disease-associated mutations, researchers employed the PMut web server. PMUT approach employs neural networks (NNs) that have been trained on a huge database of (NEMUs) neutral and pathogenic mutations of mutational hot areas identified from enormous mutations alanine scanning, and accessible genetic mutations. The mutations ranges between 0 to 1 of pathogenicity (indexes > 0.5 single mutations involved in disease causing)shown as result [43]. The name of residue and wild type position

and mutant amino acids, as well as the UniProt accession ID of TREM2, were submitted as input for this service. 2.3.3

3.2.6. PON-P2

(http://structure.bmc.lu.se/PON-P2/)

It's meant for categorizing amino acids in human proteins based on machine-learning changes. Pathogenic, neutral, and unknown amino acid changes are divided into three categories. It can quickly examine datasets on large scale. it also makes use of functional annotations and GO annotations. PON-P2 accepts nsSNP data in a number of forms. For query, submission requires substitution of amino acid and one of Ensembl or Entrez gene identifiers, UniProtKB/ accession ID[44].

3.2.7. SuSPect

http://www.sbg.bio.ic.ac.uk/suspect

SuSPect is a tool for finding the likelihood of SAVs causing disease. On the VariBench benchmarking dataset, surpasses other batch algorithms with an 82 percent balanced accuracy. SuSPect is a web tool that enables users to enter specific mutations or access a database of already calculated scores for all probable SAVs in the entire complement of human proteins [45].

3.2.8. PredictSNP

https://loschmidt.chemi.muni.cz/predictsnp/

It combines 8 tools into a classifier consensus including MAPP, PANTHER, nsSNP Analyzer, , PhD-SNP, PolyPhen-1, PolyPhen-2, SIFT and SNAP, significantly improving efficiency of prediction while also returning outcomes for all variations of mutant, proving that consent estimation is more reliable and accurate option to individual prediction tools[46].

3.2.9. Meta-SNP

http://snps.biofold.org/meta-snp/index.html

The Meta-SNP sever combines four current approaches for detecting disease-associated SNPs: PANTHER, PhD-SNP, SNAP and SIFT. The tool calculates the probability that a given nsSNP is disease-related, with values greater than 0.5 indicating that the variation is the cause of disease[47]. It's a

Forest WEKA library containing 100 nodes that was trained using 20-fold cross-validation on SV-2009 [48].

3.3. DELETERIOUS nsSNPs IDENTIFICATION USING STRUCTURE BASED PREDICTION TOOLS

Ten methods based on structural study were used to examine the protein stability variation caused by nsSNPs that were identified as deleterious by sequence-based analyses. DynaMut, SDM, CUPSAT, DUET, iStable 2.0, I-Mutant 3.0, MutPred2, MUpro, SNPs&GO3d, and Align-GVGD tools were used to estimate the effects on protein stability and dynamics by nsSNPs

3.3.1. I-Mutant3.0

http://gpcr2.biocomp.unibo.it/cgi/predictors/I-Mutant3.0/I-Mutant3.0.cgi

I-Mutant is a structure-based tool that uses SVM-based predictors to find out how single-site mutations would affect stability of a protein. Based on the DDG value (change in Gibbs free energy), it finds out whether a mutation can mainly destabilise, stabilise, or have a feeble influence on the protein. By subtraction of unfolding Gibbs free energy value of the mutant protein from the unfolding Gibbs free energy value of the wild-type protein (Kcal/mol), the DDG value is calculated. It is a server based on a SVM which work on the dataset that has been retrieved from ProTherm, a repository of empirical obsevations on protein mutations[49]. The query can be submitted in the form of a protein sequence or on a structured basis. Here, we use protein FASTA sequence for the present study[50].

3.3.2. DynaMut

http://biosig.unimelb.edu.au/dynamut/

DynaMut is a web server that implements two independent, normal mode algorithms used in analysing and visualising protein dynamics as well as checking impacts of mutations on protein stability and dynamics stability because of vibrational entropy changes as these methods are already well-established. DynaMut will provide a full suite for protein mobility and flexibility analysis and visualisation[51].

3.3.3. MUpro

http://mupro.proteomics.ics.uci.edu/

It is a SVM-based tool for finding non-synonymous SNP-induced protein stability alterations. A confidence scores in between -1 and 1 is calculated for measurement of confidence in the prediction. A score of 0 indicates that the variant reduces the stability of protein, whereas a score of >0 indicates that the variant enhances it[52].

3.3.4. MutPred2

http://mutpred.mutdb.org/

MutPred was used to predict functional and structural changes caused by amino acid substitutions. These changes were quantified as the probability of structure and function gain or loss. Furthermore, it anticipates the disease's molecular cause. The MutPred result includes a general score (g), which indicates the likelihood that the AAS is deleterious/ disease-related, as well as top 5 property scores (p), which indicate the chance that various functional and structural features are compromised. A missense mutation with a MutPred (g) score of > 0.5 might be regarded as "damaging," while (g) value of > 0.75 should be recognised a "deleterious" prediction with high confidence. Protein FASTA sequences were provided as input, along with amino acid substitutions[53].

3.3.5. DUET

http://biosig.unimelb.edu.au/duet/stability

DUET is a web server that aids –up in perusing mutations that are missense in nature in proteins by making use of integrated computational method. DUET merges the outcomes of both complementary methods (mCSM and SDM) into a prediction on consensus basis, which is generated by fusing the results of these disparate methods in an optimized predictor using SVM [54].

3.3.6. SDM

http://marid.bioc.cam.ac.uk/sdm2

The SDM method, first proposed in, uses amino acid propensities for homologous protein families derived from environment-specific substitution tables to feed a statistical potential energy function and incorporate an evolutionary idea of the constraints as such by the immediate residue environment. To

determine the free energy differences between mutant-type and wild proteins, the method examines amino acid inclination toward folded and unfolded states for wild-type and mutant proteins[55].

3.3.7. iStable 2.0

http://ncblab.nchu.edu.tw/istable2/

iStable2.0 is a web server that uses machine learning to integrate 11 sequence- and structure-based prediction algorithms and adds protein sequence information as a feature (Mutant2.0, MUpro, and iPTREE-STAB, and the structure-based tools I-Mutant2.0, PoPMuSiC, CUPSAT, AUTO-MUTE2.0, DUET, SDM, mCSM, and MAESTRO). To enhance the prediction performance of the previous version of the iStable, 3-coding modules are built for it: an Online Server Module, a Sequence Coding Module, and a Stand-alone Module. The Online Server Module is related to retrieving prediction tool results from a web server. The Stand-alone Module is responsible for delivering orders to stand-alone prediction tools and receiving their result and the sequence Coding Module for extracting and encoding information from protein sequences[56].

3.3.8. CUPSAT

http://cupsat.tu-bs.de/

CUPSAT is a web-based tool used for predicting and analysing protein stability alterations as a result of point mutations (single amino acid mutations). ΔG, the difference in free energy of unfolding between mutant proteins and wild-type, is predicted using torsion angle and structural environment-specific atom potentials. It requires a PDB protein structure as well as the mutated residue's position. The output has information of the mutation site, its structural properties like accessibility of solvent, secondary structure, torsion angles, and alteration related as well as thorough information in protein stability for 19 distinct amino acid substitutions. It also looks at how well the mutated AA adapt to the torsion angles that are detected[57].

3.3.9. Align-GVGD

http://agvgd.hci.utah.edu/

It is a web tool that utilises the physicochemical properties of AA and MSA of proteins to predict where missense mutations in genes associated lie on a continuum from harmful enriched to neutral enriched. Align-GVGD is a MSA and genuine simultaneous multiple comparisons expansion of the original

Grantham difference. Users can either provide their own MSA of proteins or use one provided by align gvgd (in FASTA format)[58].

3.3.10. SNPs&GO^{3d}

https://snps.biofold.org/snps-and-go/snps-and-go-3d.html

SNPs&GO is a strategy that uses protein functional annotation to find detrimental SAPs. The server uses Support Vector Machines (SVM), and the sequence and/or three-dimensional structure of a protein, as well as a collection of target variation and functional Gene Ontology (GO) keywords, are all inputs for a specific protein. The server's output shows the probability of each protein variant being linked to human illnesses. Findings show the server comprises two primary components: structure-based SNPs&GO3d programs and latest versions of the sequence based SNPs&GO (recently ranked as one of the top algorithms for predicting harmful SAPs). A large set of annotated variants taken from the SwissVar database is used to evaluate sequence and structure-based methods[59].

3.4. TREM2 INTERACTION ANALYSIS

Interactions of TREM2 were analysed using the STRING database, which has a predicted protein-protein signaling a wide collection of known interactions. This approach is important in the circumstances of the current study since mutations in TREM2 will undoubtedly influence the interaction network and functions the network is involved in[60].

3.5. IDENTIFICATION OF CONSERVED REGIONS OF TREM2

TREM2's evolutionarily conserved residues were identified using the ConSurf server[61]. The Bayesian method was used to generate the conservation score using the structure or protein sequence as an input query. The conservation score ranges from 1 to 9, with 1 signifying a low conservation score and 9 indicating a high conservation score. MSA was performed with the help of Clustal Omega, MUSCLE, T-Coffee and MAFFT for different species of TREM2 sequence.

Out of 26 ns-SNPs, six (Y38C, Y38H, W50S, C51G, D104G, and V124D) were found to be affected by all the tools, which were used to conduct simulations and structural analyses.

3.6. MOLECULAR DYNAMICS SIMULATIONS

Molecular dynamics simulation (MDS) analysis is an effecient tool for exposing the consequences of mutations on proteins, as proven by several comparable studies[62].GROMACS version 2018.5.0 was

used to do molecular dynamics simulations on the 3.1 Å X-ray crystal structure of TREM2 (5ELI)[63]. The MDS protocol used in this investigation is similar to that described in earlier works[64]–[67].

3.6.1. Trajectory analysis

Various GROMACS analytic tools were used to do trajectory analysis. The gmxrmsf and gmxrms programmes were used for the computation of root-mean-square deviation (RMSD) and root-mean square fluctuations (RMSF) of wild-type and mutant proteins, respectively. Furthermore, the radius of gyration (Rg), solvent accessible surface area (SASA), and intermolecular hydrogen bonding between the molecules of the provided protein were investigated using the gmxsasa, gmxhbond and gmx gyrate tools. The do_dssp utility was used to do secondary structure analysis. Gmxcovar and Gmxanaeig were also utilized to perform principal component analysis (PCA) on the complicated components to investigate their slow and functional motions[68], [69]. Finally, Visual Molecular Dynamics[70]and Chimera[71]were used to examine the trajectories. The graphs were created and visualized using the GRACE and Origin applications. (http://plasma-gate.weizmann.ac.il/Grace)

CHAPTER:4. RESULT AND THEIR DISCUSSIONS

4.1. DISTRIBUTION OF TREM2 SNPs

Total 2519 SNPs of TREM2 were retrieved from dbSNP. Among which 242 were missense SNPs, 94 were synonymous SNP, 20 were intron SNP, 2 were initiator codon variant SNP and 1 was inframe deletion SNP. The present research has considered only nsSNPs. The SNPs of TREM2 that are reported in dbSNP are distributed as listed.

4.2. SCREENING OF HIGH-RISK nsSNPs

Current research was having major focus on the nsSNPs and their regulatory function in AD and Neuroinflammation, so only nsSNPs were studied and their functional and structural effect was predicted for TREM2 protein. To find significant nsSNPs for TREM2, we utilised a total of 242 nsSNP datasets. All nsSNPs were investigated to see how they affected the protein's structure and function.

Utilizing these sequence-based prediction tools, 26 nsSNPs were usually found to be deleterious, probably damaging, or disease-associated out of 242 nsSNPs. The sequence-based tools PROVEAN, PMut, PolyPhen2, Predict SNP, Meta-SNP, PON-P2, PhD-SNP, SNAP2, and SusPect were used for sequence-based SNP prediction analysis.

These 26 nsSNPs were chosen for additional investigation since they were all expected to be important mutations by all tools. The structure-based tools DynaMut, PON-P2, SDM, Align GVGD, SNP&GO3D, MUPro, DUET, iStable, I-Mutant, and CUPSAT were used to further analyse the 26 nsSNPs selected from **Table 1**. By structure-based analysis, we predict 9 nsSNPs which is further analysed by ConSurf for conserved nsSNPs.

A total of 18-19 computational methods were utilised in this work to filter out important SNPs to get reliable findings and eliminate the chance of pseudo-positives. On combining the findings of all tools, five nsSNPs (Y38C, Y38H, W50S, D104G, V124D, C51G) were found most deleterious. These six nsSNPs rs797044603(Y38C), rs937164872(Y38H), rs549402254(W50S), rs1409131974(D104G), rs1765556646(V124D) and rs1582079884 (C51G) were selected for further MDS analysis.

Table 1: Resultant list of nsSNPs prediction using all by 9 sequence based tools.

SNP ID	MUTA-TION	Predict-SNP	PolyPhen-2	PMut	Meta-SNP	PhD-SNP	PROVEAN	PON-P2	SNAP2	SusPect
rs2234253	T96K	Deleterious	Probably damaging	Disease	Disease	Disease	Deleterious	Pathogenic	Effect	Disease
rs28937876	K186N	Deleterious	Probably damaging	Disease	Disease	Disease	Deleterious	Pathogenic	Effect	Disease
rs121908402	V126G	Deleterious	Probably damaging	Disease	Disease	Disease	Deleterious	Pathogenic	Effect	Disease
rs753372932	N68K	Deleterious	Probably damaging	Disease	Disease	Disease	Deleterious	Pathogenic	Effect	Disease
rs797044603	Y38C	Deleterious	Probably damaging	Disease	Disease	Disease	Deleterious	Pathogenic	Effect	Disease
rs369181900	C51Y	Deleterious	Probably damaging	Disease	Disease	Disease	Deleterious	Pathogenic	Effect	Disease
rs374851046	R52H	Deleterious	Probably damaging	Disease	Disease	Disease	Deleterious	Pathogenic	Effect	Disease
rs549402254	W50S	Deleterious	Probably damaging	Disease	Disease	Disease	Deleterious	Pathogenic	Effect	Disease
rs746216516	S31F	Deleterious	Probably damaging	Disease	Disease	Disease	Deleterious	Pathogenic	Effect	Disease
rs747377941	S65G	Deleterious	Probably damaging	Disease	Disease	Disease	Deleterious	Pathogenic	Effect	Disease
rs749358844	R52C	Deleterious	Probably damaging	Disease	Disease	Disease	Deleterious	Pathogenic	Effect	Disease
rs753325601	R47C	Deleterious	Probably damaging	Disease	Disease	Disease	Deleterious	Pathogenic	Effect	Disease
rs760511786	T22I	Deleterious	Probably damaging	Disease	Disease	Disease	Deleterious	Pathogenic	Effect	Disease
rs764489173	S158R	Deleterious	Probably damaging	Disease	Disease	Disease	Deleterious	Pathogenic	Effect	Disease
rs774054773	G80R	Deleterious	Probably damaging	Disease	Disease	Disease	Deleterious	Pathogenic	Effect	Disease
rs779888024	Y108C	Deleterious	Probably damaging	Disease	Disease	Disease	Deleterious	Pathogenic	Effect	Disease

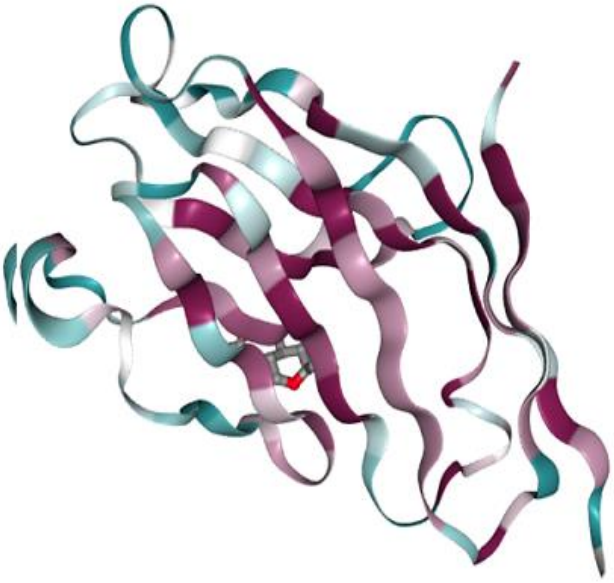
rs937164872	Y38H	Deleterious	Probably damaging	Disease	Disease	Disease	Deleterious	Pathogenic	Effect	Disease
rs1204617414	L93R	Deleterious	Probably damaging	Disease	Disease	Disease	Deleterious	Pathogenic	Effect	Disease
rs1311610930	L97P	Deleterious	Probably damaging	Disease	Disease	Disease	Deleterious	Pathogenic	Effect	Disease
rs1409131974	D104G	Deleterious	Probably damaging	Disease	Disease	Disease	Deleterious	Pathogenic	Effect	Disease
rs1582079884	C51G	Deleterious	Probably damaging	Disease	Disease	Disease	Deleterious	Pathogenic	Effect	Disease
rs1582080006	V34G	Deleterious	Probably damaging	Disease	Disease	Disease	Deleterious	Pathogenic	Effect	Disease
rs1765556646	V124D	Deleterious	Probably damaging	Disease	Disease	Disease	Deleterious	Pathogenic	Effect	Disease
rs1765566438	R46G	Deleterious	Probably damaging	Disease	Disease	Disease	Deleterious	Pathogenic	Effect	Disease
rs61608747	T96M	Deleterious	Probably damaging	Disease	Disease	Disease	Deleterious	Pathogenic	Effect	Disease
rs75932628	R47H	Deleterious	Probably damaging	Disease	Disease	Disease	Deleterious	Pathogenic	Effect	Disease

Table 2: Resultant list of nsSNPs prediction using all 10 structure-based tools.

SNP ID	MUTA-TION	I-Mutant	DynaMut	MUpro	iStatble 2.0	CUPSAT	MutPred2	DUET	Align GVGD	SDM	SNP& GO^{3D}
rs121908402	**V126G**	Decrease	Destabilising	Decrease	Decrease	Destabilising	Deleterious	Destabilising	Most Likely	Reduced stability	Disease
rs797044603	**Y38C**	Decrease	Destabilising	Decrease	Decrease	Destabilising	Deleterious	Destabilising	Most Likely	Reduced stability	Disease
rs549402254	**W50S**	Decrease	Destabilising	Decrease	Decrease	Destabilising	Deleterious	Destabilising	Most Likely	Reduced stability	Disease
rs1204617414	**L93R**	Decrease	Destabilising	Decrease	Decrease	Destabilising	Deleterious	Destabilising	Most Likely	Reduced stability	Disease
rs1311610930	**L97P**	Decrease	Destabilising	Decrease	Decrease	Destabilising	Deleterious	Destabilising	Most Likely	Reduced stability	Disease
rs1409131974	**D104G**	Decrease	Destabilising	Decrease	Decrease	Destabilising	Deleterious	Destabilising	Most Likely	Reduced stability	Disease
rs1582080006	**V34G**	Decrease	Destabilising	Decrease	Decrease	Destabilising	Deleterious	Destabilising	Most Likely	Reduced stability	Disease
rs1765556646	**V124D**	Decrease	Destabilising	Decrease	Decrease	Destabilising	Deleterious	Destabilising	Most Likely	Reduced stability	Disease
rs1765566438	**R46G**	Decrease	Destabilising	Decrease	Decrease	Destabilising	Deleterious	Destabilising	Most Likely	Reduced stability	Disease

4.3. CONSERVATION ANALYSIS OF TREM2

The conservation level residue provides an estimate of the structural and functional impact that deleterious mutations to the protein potentially cause. A deleterious mutation having highly conserved score is always damaging in nature. The structure of protein (5ELI) was provided as an query for ConSurf, and the Bayesian approach was utilised to calculate the results. Residues can have conservation scores ranging from 1 to 9. A residue with a score of "1" is highly variable, whereas a residue with a value of "9" is highly conserved. The ConSurf results described that six mutations Y38C, Y38H, W50S, D104G, V124D, and C51G, each with a conservation score of 9 at highly conserved residues. **Figure 1** shows the structural representation of the conservation results for TREM 2 chain A. Multiple sequence alignment (MSA) was used to confirm the ConSurf server's findings, and it discovered that Y38C, Y38H, W50S, C51G, D104G, and V124D were highly conserved residues using Clustal Omega, T- COFFEE, MUSCLE, and MAFFT alignment tools utilised in this work.



SR No.	Variant ID	Mutation	Residue	Color	SCORE
1	rs797044603	Y38C	Y	9	-1.368
2	rs549402254	W50S	W	9	-1.133
3	rs1409131974	D104G	D	9	-1.41
4	rs1765556646	V124D	V	9	-1.142
5	rs937164872	Y38H	Y	9	-1.368
6	rs1582079884	C51G	C	9	-1.279

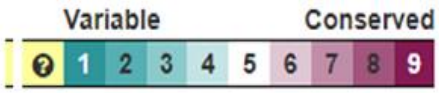


Figure. 1. All 6 potent SNPs were found in the most conserved region of the protein when measured at its TREM2 structure using the ConSurf server.

4.4. INTERACTION ANALYSIS OF TREM 2

Figure 2 shows the TREM2 interaction network retrieved from the STRING database. The input name was "TREM2," and the organism was "Homo sapiens." Proteins are represented through nodes and their respective interactions through edges. These proteins mutually add to imparted capacities to TREM2. These scores demonstrate certainty, in the other words how much probable an association is to be valid, determined by STRING based on the accessible proof. The functions will be impacted when the 6 deleterious nsSNPs have an influence on TREM2's interaction with the genes listed in **Table 3**.

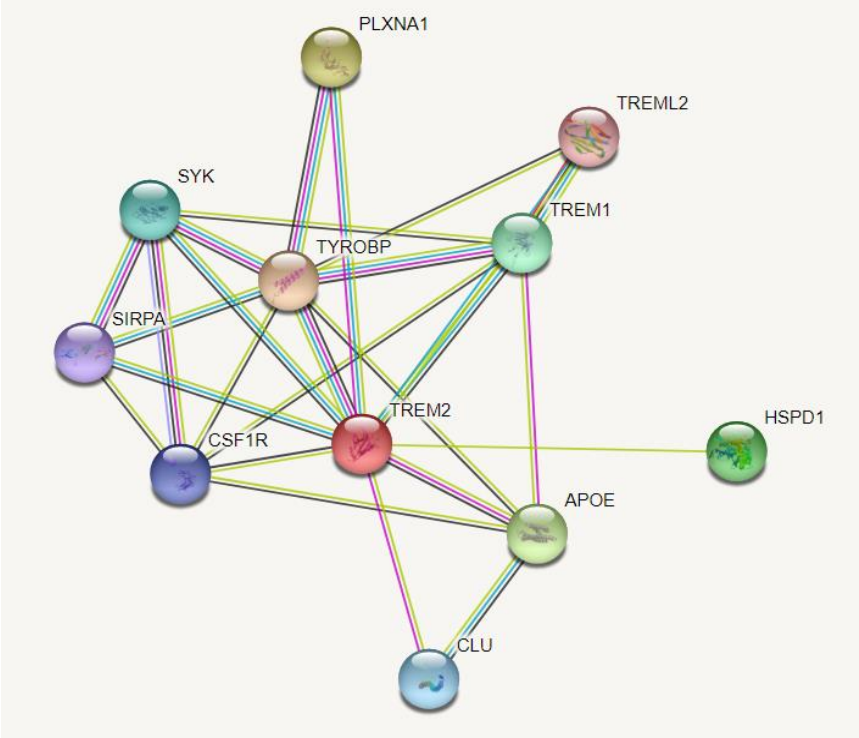


Figure:2 TREM2 interaction network obtained from the STRING database. Besides the TREM family, a few other proteins were also observed interacting.

Table. 3.

Interaction detail of the Predicted TREM2 interacting partner using STRING database.

Functional Partner	Score	Information
TYROBP	0.999	TYRO protein tyrosine kinase-binding protein; Non-covalent binding to CD300 family activating receptors. Cellular activation occurs when CD300-TYROBP complexes crosslink.
PLXNA1	0.996	Plexin-A1; SEMA3A, SEMA3C, SEMA3F and SEMA6D coreceptors Required for class 3 semaphorin signaling and subsequent cytoskeleton remodelling. Axon guidance, invasive growth, and cell migration are all functions of this protein. The plexin controls the complex's affinity for certain semaphorins, and its cytoplasmic domain is needed for the activation of cytoplasmic signaling processes.
APOE	0.995	Apolipoprotein The binding, internalisation, and degradation of lipoprotein particles are all mediated by this protein. It can act as a ligand for the low density lipoprotein (apo B/E) and the apo-E receptors found in hepatic tissues (chylomicron remnant).
HSPD	0.975	Heat shock protein family d (hsp60) member 1 Chaperonin has been linked to macromolecular assembly and imports mitochondrial protein. In the mitochondrial matrix, it may also prevent misfolding and promote refolding and appropriate assembly of unfolded polypeptides formed under stress circumstances. These chaperonins' functional units are heptameric rings of the major subunit Hsp60 that operate as a back-to-back double ring.
TREM1	0.944	Triggering receptor expressed on myeloid cells 1 Neutrophils and monocytes play a major role in inflammatory responses. Increases the synthesis of proinflammatory chemokines and cytokines, as well as the expression of cell activation markers on the surface of cells.

SYK	0.932	Spleen associated tyrosine kinase Signaling downstream of a number of transmembrane receptors, including conventional immunoreceptors like the B-cell receptor, is regulated by this non-receptor tyrosine kinase (BCR). This protein regulates biological processes such as Innate and adaptive immunity, cell adhesion, osteoclast initiation, platelet activation, and vascular function.
CLU	0.927	Clusterin Isoform 1 acts as an extracellular chaperone, preventing non-native proteins from aggregating. Prevents blood plasma protein aggregation caused by stress. APP, B2M, SNCA, CALCA, CSN3, APOC2, and aggregation-prone LYZ variations are all inhibited from forming amyloid fibrils (in vitro). Maintains partly unfolded proteins in a condition that allows some other chaperones, such as HSPA8/HSC70, to refold them.
CSF1R	0.914	Macrophage colony-stimulating factor 1 receptor Tyrosine-protein kinase that affects the survival, proliferation, and differentiation of hematopoietic progenitor cells, notably mononuclear phagocytes like macrophages and monocytes, by acting as a cell-surface receptor for CSF1 and IL34. Promotes the synthesis in response to IL34 and CSF1 of proinflammatory chemokines, and hence plays a major role in inflammatory conditions and innate immunity.
SIRPA	0.900	Tyrosine-protein phosphatase non-receptor type substrate 1 CD47 cell surface receptor that is immunoglobulin-like in nature. PTPN6, PTPN11, and other binding partners are translocated from the cytosol to the plasma membrane by this protein, which acts as a docking protein. Cerebellar neuron adhesion, neurite development, and glial cell attachment are all aided when this protein is present. It may be important for synaptogenesis and synaptic function during intracellular signaling (By similarity). Negatively regulates cell adhesion, growth factors, and insulin-stimulated receptor tyrosine kinase-coupled cellular responses.
TREML2	0.850	TREM-like transcript 2 protein The immune responses like innate and adaptive may be aided by this cell surface receptor. Interaction with CD276 on T-cells promotes T-cell activation and acts as a counter-receptor for CD276.

4.5. MOLECULAR DYNAMICS SIMULATIONS

MDS were used to investigate the deviation in conformation between the WT and mutant proteins. We have carried out 200 ns MDS and then various structural parameters were analyzed and described below in detail (Wild-Type Y38C, Y38H, W50S, D104G, V124D, and C51G). Then we looked at things like RMSD, RMSF, hydrogen bonds, and PCA.

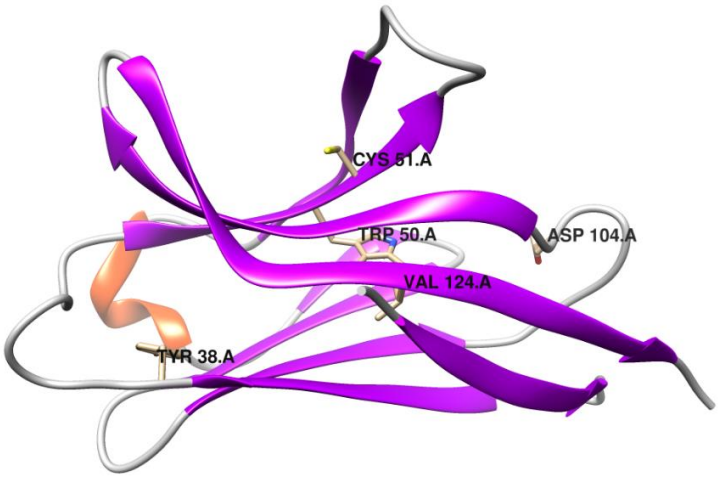


Figure.3. WT residues and mutated AA residues positions of highly-risked nsSNPs analysed using Tools.

4.5.1. RMSD (Relative Mean Square Deviation)

The RMSD of all the Cα atoms in the original structure was computed. The variations obtained during the simulation can be used to determine the stability of the protein in relation to its conformation. Protein structures with less variations are more stable. The RMSD for all the systems were calculated and shown in **Figure**. Average RMSD for WT, Y38C, Y38H, W50S, C51G, D104G, and V124D were 0.12, 0.13, 0.13, 0.12, 0.14, 0.11 and 0.14 nm respectively. The RMSD result demonstrates that Y38C, Y38H, C51G, and V124D are unstable because they are showing higher RMSD values. The D104G showed less RMSD value than the WT hence it represents that it is more stable, whereas the W50S is showing equal RMSD value as like WT therefore it represents that the mutation is not inducing instability in the protein. From the overall RMSD analysis we have found that Y38C, Y38H, C51G, and V124D are inducing the instability in the protein.

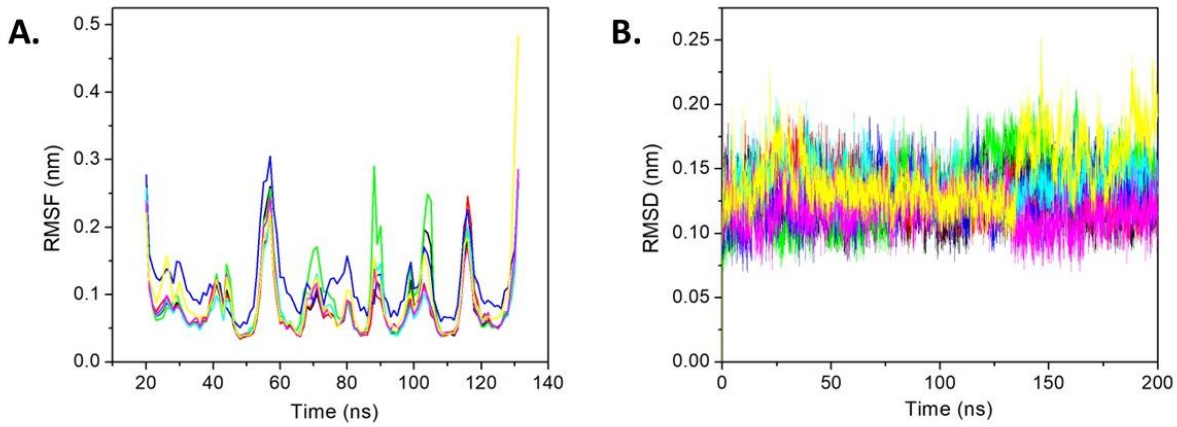


Figure.4. MDS analysis of WT and mutant proteins (A) RMSF (B) RMSD

Color scheme for panel: WT(Black), Y38C(Red), Y38H(Green), W50S(Blue), C51G(Cyan), D104G(Magenta), and V124D(Yellow)

4.5.2. RMSF (Root Mean Square Fluctuation)

The RMSF of Cα atoms was studied for both systems for a comprehensive investigation of the mutational influence on the fluctuation of the overall protein as well as particular residues. RMSF shows how mutations affect the wild-type and mutant proteins residue-by-residue mobility. The RMSF for all the systems were calculated and shown in **Figure 4(A)**. Average RMSF for WT(Black), Y38C(Red), Y38H(Green), W50S(Blue), C51G(Cyan), D104G(Magenta), and V124D(Yellow) were 0.08, 0.08, 0.09, 0.11, 0.08, 0.08 and 0.09 nm respectively. The RMSD result demonstrates that Y38H, W50S, and V124D are unstable because they are showing higher RMSF values. Y38C, C51G, and D104G are showing equal RMSF value as like WT therefore it represents that the mutation is not inducing instability in the protein. Whereas Y38H, W50S, and V124D induce stability of the protein.

4.5.3. Structural stability

The Rg, SASA, and H-bond analysis were used to examine the conformational geometry of the WT and mutant proteins.

Firstly, we have analysed the Rg value and plotted it in **Figure 5 (B)**. We have seen the average Rg value for WT, Y38C, Y38H, W50S, C51G, D104G, and V124D was 1.41, 1.40, 1.41, 1.42, 1.35, 1.41, and 1.41 nm. The Rg result represents that the mutations are not altering the compactness of the protein except C51G. The C51G showed a very low Rg value as compared to all other mutants. Other mutations showed the very similar type of Rg value therefore from here we have concluded that mutations are not able to disrupt the compactness of the protein.

The we have analysed the H-Bond value and plotted it in **Figure 5 (A).** The average H-Bond value for WT, Y38C, Y38H, W50S, C51G, D104G, and V124D was 80.75, 81.36, 76.71, 77.36, 79.69, 78.14, and 80.41 nm^2, respectively.

After H-Bond analysis, we have analysed the SASA value and plotted it in **Figure 5(C)**. The average SASA value for WT, Y38C, Y38H, W50S, C51G, D104G, and V124D was 69.9, 70.54, 70.02, 70.26, 69.77, 70.39, and 70.12 nm^2 respectively. The SASA value also showed Rg like pattern. Here also we have observed that these mutations are not able to alter the solvent accessible area while they can only alter the catalytic site.

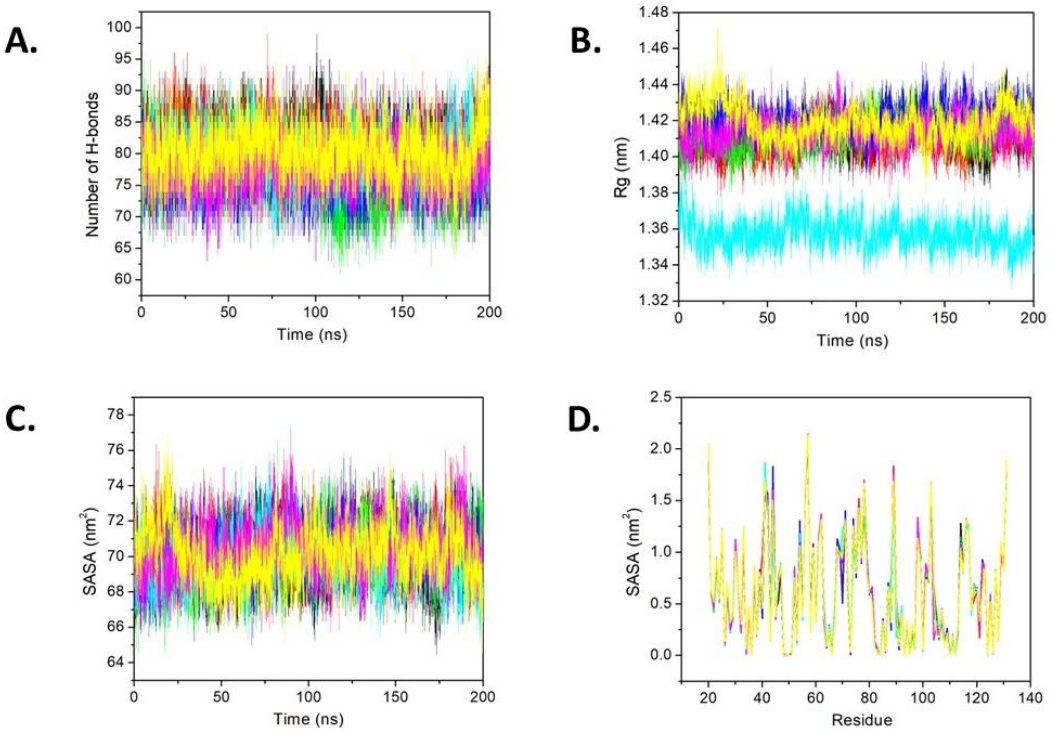


Figure. 5. MDS analysis of wild-type and mutant proteins (A) Hydrogen bonds (B) Radius of gyration (C) SASA. Color scheme for panel: WT(Black), Y38C(Red), Y38H(Green), W50S(Blue), C51G(Cyan), D104G(Magenta), and V124D(Yellow)

4.5.4. PCA (Principal Component Analysis)

PCA is a valuable tool for analysing any system's dynamic mechanical characteristics. (70) PCA was used to estimate the massive aggregate movements of the wild-type and mutant TREM2 to validate our MDS findings. The eigenvectors of a covariance matrix are called principal components (PCs). The PCA for 50 eigenvectors were calculated and shown in **Figure 6(A)**. We have considered the first 5

eigenvectors and calculated the % wise correlated motions. We have seen that WT, Y38C, Y38H, W50S, C51G, D104G, and V124D showed 50%, 52.38%, 56.19%, 49.42%, 44.57%, 45.78% and 61.90% correlated motions respectively. The PCA result clearly indicates that mutations are inducing the correlated motions because we have not seen a similar value like the WT. The C51G, D104G, and W50S showed less value which represents that the mutations are inducing higher stability in the structure while Y38C, Y38H, and V124D are showing higher values which represent those mutations are inducing instability in the protein. All the mutations are causing the alteration in the protein therefore we have concluded that these mutations are disrupting the proper binding site hence the protein will lose its function.

We can see from **Figure 6(C)** the PCA analysis that the first few eigenvectors are crucial in representing the general dynamics, therefore we've chosen only the first two eigenvectors and displayed them against each other. The Figure clearly showed that WT showed a very dense and stable cluster while mutations are disrupting the structure therefore, we have seen an unstable cluster for all the mutants. This result is also showing that mutations are causing instability in the WT protein.

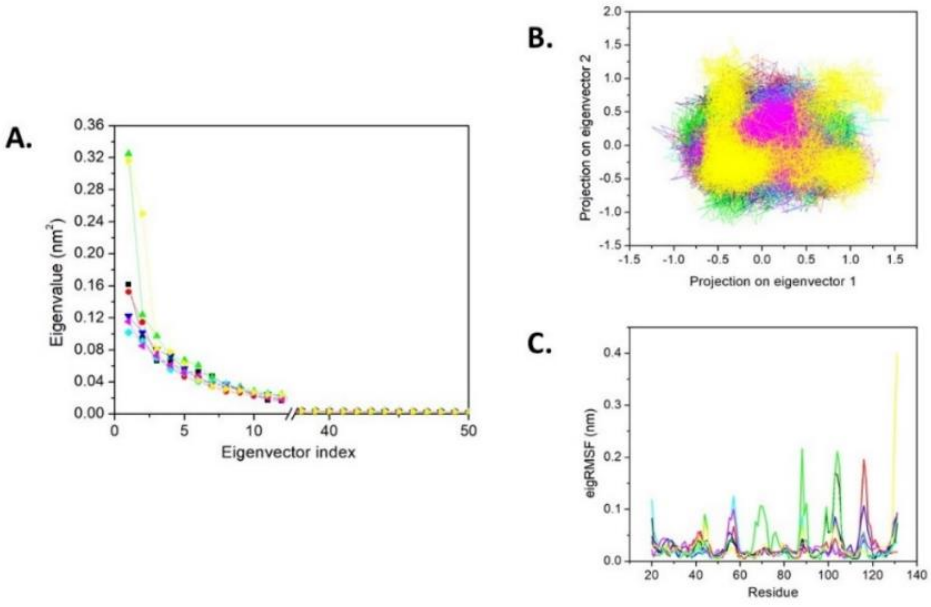


Figure 6. PCA. (A) The graphical plot between eigenvalues and eigenvector index. The first 50 eigenvectors of WT and mutant structures were plotted (B) Projection on eigenvector 1 vs Projection on eigenvector 2.
(C) Residue EigRMSF value.

CONCLUSION

AD is a chronic neurological disease that worsens with time. It is permeated by the abnormal accumulation of senile plaques and neurofibrillary tangles. These accumulations trigger the neuroinflammation which causes neuron damage. TREM2 plays a key role in the neuroinflammation process while a mutation in TREM2 alters its functions. Therefore, we have analysed the TREM2 mutations using sequence and structural related tools. We have retrieved 242 nsSNPs from the dbSNP database and then these were analysed by 9 sequence oriented tools. From this analysis, we have found 26 mutations that were found damaging in nature. Hence, these mutations were selected and used for the structure related analysis. The structure level analysis revealed 9 mutations which can affect the TREM2 conformations. Finally, to evaluate the structure level changes we have selected 6 mutations and employed them for 100 ns MDS. The MDS analysis revealed that all of the mutations are altering the proper geometry of the TREM2 protein structure therefore, from this analysis we have proposed that Y38C, Y38H, W50S, C51G, D104G, and V124D are the mutations which can be further explored by using the *in-vitro* and *in-vivo* investigations.

CHAPTER:6. REFERENCES

[1] D. Kempuraj*et al.*, "Neuroinflammation Induces Neurodegeneration," *J. Neurol. Neurosurg. Spine*, vol. 1, no. 1, p. 1003, 2016.

[2] R. B. Maccioni, L. E. Rojo, J. A. Fernández, and R. O. Kuljis, "The role of neuroimmunomodulation in Alzheimer's disease," *Ann. N. Y. Acad. Sci.*, vol. 1153, pp. 240–246, Feb. 2009, doi: 10.1111/j.1749-6632.2008.03972.x.

[3] M. D. Weingarten, A. H. Lockwood, S. Y. Hwo, and M. W. Kirschner, "A protein factor essential for microtubule assembly," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 72, no. 5, pp. 1858–1862, May 1975, doi: 10.1073/pnas.72.5.1858.

[4] D. N. Drechsel, A. A. Hyman, M. H. Cobb, and M. W. Kirschner, "Modulation of the dynamic instability of tubulin assembly by the microtubule-associated protein tau," *Mol. Biol. Cell*, vol. 3, no. 10, pp. 1141–1154, Oct. 1992, doi: 10.1091/mbc.3.10.1141.

[5] M. Citron, "Alzheimer's disease: strategies for disease modification," *Nat. Rev. Drug Discov.*, vol. 9, no. 5, pp. 387–398, May 2010, doi: 10.1038/nrd2896.

[6] M. Violet *et al.*, "A major role for Tau in neuronal DNA and RNA protection in vivo under physiological and hyperthermic conditions," *Front. Cell. Neurosci.*, vol. 8, 2014, Accessed: May 26, 2022. [Online]. Available: https://www.frontiersin.org/article/10.3389/fncel.2014.00084

[7] T. J. Malia *et al.*, "Epitope mapping and structural basis for the recognition of phosphorylated tau by the anti-tau antibody AT8," *Proteins*, vol. 84, no. 4, pp. 427–434, Apr. 2016, doi: 10.1002/prot.24988.

[8] T. Bussière*et al.*, "Phosphorylated serine422 on tau proteins is a pathological epitope found in several diseases with neurofibrillary degeneration," *Acta Neuropathol. (Berl.)*, vol. 97, no. 3, pp. 221–230, Mar. 1999, doi: 10.1007/s004010050978.

[9] L. Wang *et al.*, "Expression of Tau40 induces activation of cultured rat microglial cells," *PloS One*, vol. 8, no. 10, p. e76057, Jan. 2013, doi: 10.1371/journal.pone.0076057.

[10] H. Braak and E. Braak, "Neuropathological stageing of Alzheimer-related changes," *Acta Neuropathol. (Berl.)*, vol. 82, no. 4, pp. 239–259, 1991, doi: 10.1007/BF00308809.

[11] N. Bazan, A. Halabi, M. Ertel, and N. Petasis, "Chapter 34 – Neuroinflammation," 2012. doi: 10.1016/B978-0-12-374947-5.00034-1.

[12] C. K. Glass, K. Saijo, B. Winner, M. C. Marchetto, and F. H. Gage, "Mechanisms Underlying Inflammation in Neurodegeneration," *Cell*, vol. 140, no. 6, pp. 918–934, Mar. 2010, doi: 10.1016/j.cell.2010.02.016.

[13] J. Tr, von S. Ve, and L. Ge, "TREM2 in Neurodegenerative Diseases," *Mol. Neurodegener.*, vol. 12, no. 1, Aug. 2017, doi: 10.1186/s13024-017-0197-5.

[14] Y. Wang *et al.*, "TREM2 lipid sensing sustains the microglial response in an Alzheimer's disease model," *Cell*, vol. 160, no. 6, pp. 1061–1071, Mar. 2015, doi: 10.1016/j.cell.2015.01.049.

[15] P. Wunderlich, K. Glebov, N. Kemmerling, N. T. Tien, H. Neumann, and J. Walter, "Sequential proteolytic processing of the triggering receptor expressed on myeloid cells-2 (TREM2) protein by ectodomain shedding and γ-secretase-dependent intramembranous cleavage," *J. Biol. Chem.*, vol. 288, no. 46, pp. 33027–33036, Nov. 2013, doi: 10.1074/jbc.m113.517540.

[16] K. Schlepckow*et al.*, "An Alzheimer-associated TREM2 variant occurs at the ADAM cleavage site and affects shedding and phagocytic function," *EMBO Mol. Med.*, vol. 9, no. 10, pp. 1356–1365, Oct. 2017, doi: 10.15252/emmm.201707672.

[17] J. C. Lambert *et al.*, "Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease," *Nat. Genet.*, vol. 45, no. 12, pp. 1452–1458, Dec. 2013, doi: 10.1038/ng.2802.

[18] L. Zhong *et al.*, "DAP12 stabilizes the C-terminal fragment of the triggering receptor expressed on myeloid cells-2 (TREM2) and protects against LPS-induced pro-inflammatory response," *J. Biol. Chem.*, vol. 290, no. 25, Art. no. 25, 2015, doi: 10.1074/jbc.M115.645986.

[19] W. Zhu, H. Zheng, X. Shao, W. Wang, Q. Yao, and Z. Li, "Excitotoxicity of TNFα derived from KA activated microglia on hippocampal neurons in vitro and in vivo," *J. Neurochem.*, vol. 114, no. 2, pp. 386–396, 2010, doi: 10.1111/j.1471-4159.2010.06763.x.

[20] X. Zhu *et al.*, "TL-2 attenuates β-amyloid induced neuronal apoptosis through the AKT/GSK-3β/β-catenin pathway.," *Int. J. Neuropsychopharmacol.*, 2014, doi: 10.1017/S1461145714000315.

[21] N. A-Gonzalez *et al.*, "Apoptotic cells promote their own clearance and immune tolerance through activation of the nuclear receptor LXR," *Immunity*, vol. 31, no. 2, pp. 245–258, Aug. 2009, doi: 10.1016/j.immuni.2009.06.018.

[22] A. Miyashita *et al.*, "Lack of genetic association between TREM2 and late-onset Alzheimer's disease in a Japanese population," *J. Alzheimers Dis. JAD*, vol. 41, no. 4, pp. 1031–1038, 2014, doi: 10.3233/JAD-140225.

[23] S. J. Chung *et al.*, "Exome array study did not identify novel variants in Alzheimer's disease," *Neurobiol. Aging*, vol. 35, no. 8, p. 1958.e13-1958.e14, Aug. 2014, doi: 10.1016/j.neurobiolaging.2014.03.007.

[24] B. Hooli*et al.*, "The rare TREM2 R47H variant exerts only a modest effect on Alzheimer disease risk," *Neurology*, vol. 83, Sep. 2014, doi: 10.1212/WNL.0000000000000855.

[25] L. Bertram, A. R. Parrado, and R. E. Tanzi, "TREM2 and neurodegenerative disease," *N. Engl. J. Med.*, vol. 369, no. 16, p. 1565, Oct. 2013, doi: 10.1056/NEJMc1306509.

[26] O. Korvatska*et al.*, "R47H Variant of TREM2 Associated With Alzheimer Disease in a Large Late-Onset Family: Clinical, Genetic, and Neuropathological Study," *JAMA Neurol.*, vol. 72, no. 8, pp. 920–927, Aug. 2015, doi: 10.1001/jamaneurol.2015.0979.

[27] T. Jiang *et al.*, "A rare coding variant in TREM2 increases risk for Alzheimer's disease in Han Chinese," *Neurobiol. Aging*, vol. 42, p. 217.e1-217.e3, Jun. 2016, doi: 10.1016/j.neurobiolaging.2016.02.023.

[28] S. C. Jin*et al.*, "TREM2 is associated with increased risk for Alzheimer's disease in African Americans," *Mol. Neurodegener.*, vol. 10, p. 19, Apr. 2015, doi: 10.1186/s13024-015-0016-9.

[29] B. A. Benitez *et al.*, "Missense variant in TREML2 protects against Alzheimer's disease," *Neurobiol. Aging*, vol. 35, no. 6, p. 1510.e19–26, Jun. 2014, doi: 10.1016/j.neurobiolaging.2013.12.010.

[30] M. M. Carrasquillo *et al.*, "A candidate regulatory variant at the TREM gene cluster associates with decreased Alzheimer's disease risk and increased TREML1 and TREM2 brain gene expression," *Alzheimers Dement.*, vol. 13, no. 6, pp. 663–673, Jun. 2017, doi: 10.1016/j.jalz.2016.10.005.

[31] C. F. Slattery *et al.*, "R47H TREM2 variant increases risk of typical early-onset Alzheimer's disease but not of prion or frontotemporal dementia," *Alzheimers Dement.*, vol. 10, no. 6, pp. 602-608.e4, Nov. 2014, doi: 10.1016/j.jalz.2014.05.1751.

[32] S. L. Rosenthal *et al.*, "More evidence for association of a rare TREM2 mutation (R47H) with Alzheimer's disease risk," *Neurobiol. Aging*, vol. 36, no. 8, p. 2443.e21-2443.e26, Aug. 2015, doi: 10.1016/j.neurobiolaging.2015.04.012.

[33] P. Roussos, P. Katsel, P. Fam, W. Tan, D. P. Purohit, and V. Haroutunian, "The triggering receptor expressed on myeloid cells 2 (TREM2) is associated with enhanced inflammation,

neuropathological lesions and increased risk for Alzheimer's Dementia," *Alzheimers Dement. J. Alzheimers Assoc.*, vol. 11, no. 10, pp. 1163–1170, Oct. 2015, doi: 10.1016/j.jalz.2014.10.013.

[34] E. O. Luis *et al.*, "Frontobasalgray matter loss is associated with the TREM2 p.R47H variant," *Neurobiol. Aging*, vol. 35, no. 12, pp. 2681–2690, Dec. 2014, doi: 10.1016/j.neurobiolaging.2014.06.007.

[35] C. M. Lill*et al.*, "The role of TREM2 R47H as a risk factor for Alzheimer's disease, frontotemporal lobar degeneration, amyotrophic lateral sclerosis, and Parkinson's disease," *Alzheimers Dement. J. Alzheimers Assoc.*, vol. 11, no. 12, pp. 1407–1416, Dec. 2015, doi: 10.1016/j.jalz.2014.12.009.

[36] C. Cruchaga*et al.*, "GWAS of cerebrospinal fluid tau levels identifies risk variants for Alzheimer's disease," *Neuron*, vol. 78, no. 2, pp. 256–268, Apr. 2013, doi: 10.1016/j.neuron.2013.02.026.

[37] "TREM2 variants and risk of Alzheimer's disease: a meta-analysis | SpringerLink." https://link.springer.com/article/10.1007/s10072-015-2274-2 (accessed May 26, 2022).

[38] C. Engelman, R. Koscik, E. Jonaitis, B. Hermann, A. L. Rue, and M. Sager, "Investigation of triggering receptor expressed on myeloid cells 2 variant in the Wisconsin Registry for Alzheimer's Prevention," *Neurobiol. Aging*, 2014, doi: 10.1016/j.neurobiolaging.2013.11.013.

[39] Y. Choi and A. P. Chan, "PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels," *Bioinforma. Oxf. Engl.*, vol. 31, no. 16, pp. 2745–2747, Aug. 2015, doi: 10.1093/bioinformatics/btv195.

[40] I. A. Adzhubei*et al.*, "A method and server for predicting damaging missense mutations," *Nat. Methods*, vol. 7, no. 4, pp. 248–249, Apr. 2010, doi: 10.1038/nmeth0410-248.

[41] L. Hecht, J. Wass, L. Kelly, E. Clevenger-Firley, and C. Dunn, "SNAP-Ed Steps to Health inspires third graders to eat smart and move more.," *J. Nutr. Educ. Behav.*, 2013, doi: 10.1016/j.jneb.2013.02.007.

[42] E. Capriotti and R. Calabrese, "Predicting the insurgence of human genetic diseases associated to single point protein mutations with support vector machines and evolutionary information," *Bioinformatics*, vol. 22, pp. 2729–2734, Jan. 2006.

[43] C. Ferrer-Costa, J. Gelpí, L. Zamakola, I. Parraga, X. Cruz, and M. Orozco, "PMUT: a web-based tool for the annotation of pathological mutations on proteins," *Bioinform.*, 2005, doi: 10.1093/bioinformatics/bti486.

[44] A. Niroula, S. Urolagin, and M. Vihinen, “PON-P2: Prediction Method for Fast and Reliable Identification of Harmful Variants,” *PLOS ONE*, vol. 10, no. 2, p. e0117380, Feb. 2015, doi: 10.1371/journal.pone.0117380.

[45] C. M. Yates, I. Filippis, L. A. Kelley, and M. J. E. Sternberg, “SuSPect: Enhanced Prediction of Single Amino Acid Variant (SAV) Phenotype Using Network Features,” *J. Mol. Biol.*, vol. 426, no. 14, pp. 2692–2701, Jul. 2014, doi: 10.1016/j.jmb.2014.04.026.

[46] J. Bendl*et al.*, “PredictSNP: robust and accurate consensus classifier for prediction of disease-related mutations,” *PLoSComput. Biol.*, vol. 10, no. 1, p. e1003440, Jan. 2014, doi: 10.1371/journal.pcbi.1003440.

[47] E. Capriotti, R. B. Altman, and Y. Bromberg, “Collective judgment predicts disease-associated single nucleotide variants,” *BMC Genomics*, vol. 14, no. 3, p. S2, May 2013, doi: 10.1186/1471-2164-14-S3-S2.

[48] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, “The WEKA data mining software: an update,” *ACM SIGKDD Explor. Newsl.*, vol. 11, no. 1, pp. 10–18, Nov. 2009, doi: 10.1145/1656274.1656278.

[49] K. A. Bava, M. M. Gromiha, H. Uedaira, K. Kitajima, and A. Sarai, “ProTherm, version 4.0: thermodynamic database for proteins and mutants,” *Nucleic Acids Res.*, vol. 32, no. suppl_1, pp. D120–D121, Jan. 2004, doi: 10.1093/nar/gkh082.

[50] E. Capriotti, P. Fariselli, and R. Casadio, “I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure,” *Nucleic Acids Res.*, vol. 33, no. Web Server issue, pp. W306-310, Jul. 2005, doi: 10.1093/nar/gki375.

[51] C. H. Rodrigues, D. E. Pires, and D. B. Ascher, “DynaMut: predicting the impact of mutations on protein conformation, flexibility and stability,” *Nucleic Acids Res.*, vol. 46, no. Web Server issue, pp. W350–W355, Jul. 2018, doi: 10.1093/nar/gky300.

[52] J. Cheng, A. Randall, and P. Baldi, “Prediction of protein stability changes for single-site mutations using support vector machines,” *Proteins*, vol. 62, no. 4, pp. 1125–1132, Mar. 2006, doi: 10.1002/prot.20810.

[53] V. Pejaver*et al.*, “Inferring the molecular and phenotypic impact of amino acid variants with MutPred2,” *Nat. Commun.*, vol. 11, no. 1, p. 5918, Nov. 2020, doi: 10.1038/s41467-020-19669-x.

[54] D. E. V. Pires, D. B. Ascher, and T. L. Blundell, "DUET: a server for predicting effects of mutations on protein stability using an integrated computational approach," *Nucleic Acids Res.*, vol. 42, no. Web Server issue, pp. W314-319, Jul. 2014, doi: 10.1093/nar/gku411.

[55] C. L. Worth, R. Preissner, and T. L. Blundell, "SDM—a server for predicting effects of mutations on protein stability and malfunction," *Nucleic Acids Res.*, vol. 39, no. Web Server issue, pp. W215–W222, Jul. 2011, doi: 10.1093/nar/gkr363.

[56] C.-W. Chen, M.-H. Lin, C.-C. Liao, H.-P. Chang, and Y.-W. Chu, "iStable 2.0: Predicting protein thermal stability changes by integrating various characteristic modules," *Comput. Struct. Biotechnol. J.*, vol. 18, pp. 622–630, 2020, doi: 10.1016/j.csbj.2020.02.021.

[57] V. Parthiban, M. M. Gromiha, and D. Schomburg, "CUPSAT: prediction of protein stability upon point mutations," *Nucleic Acids Res.*, vol. 34, no. Web Server issue, pp. W239-242, Jul. 2006, doi: 10.1093/nar/gkl190.

[58] S. V. Tavtigian*et al.*, "Comprehensive statistical study of 452 BRCA1 missense substitutions with classification of eight recurrent substitutions as neutral," *J. Med. Genet.*, vol. 43, no. 4, pp. 295–305, Apr. 2006, doi: 10.1136/jmg.2005.033878.

[59] E. Capriotti, R. Calabrese, P. Fariselli, P. L. Martelli, R. B. Altman, and R. Casadio, "WS-SNPs&GO: a web server for predicting the deleterious effect of human protein variants using functional annotation," *BMC Genomics*, vol. 14 Suppl 3, p. S6, 2013, doi: 10.1186/1471-2164-14-S3-S6.

[60] D. Szklarczyk*et al.*, "STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets," *Nucleic Acids Res.*, vol. 47, no. D1, pp. D607–D613, Jan. 2019, doi: 10.1093/nar/gky1131.

[61] F. Glaser *et al.*, "ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information," *Bioinforma. Oxf. Engl.*, vol. 19, no. 1, pp. 163–164, Jan. 2003, doi: 10.1093/bioinformatics/19.1.163.

[62] B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl, "GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation," *J. Chem. Theory Comput.*, vol. 4, no. 3, pp. 435–447, Mar. 2008, doi: 10.1021/ct700301q.

[63] S. Pronk *et al.*, "GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit," *Bioinforma. Oxf. Engl.*, vol. 29, no. 7, pp. 845–854, Apr. 2013, doi: 10.1093/bioinformatics/btt055.

[64] A. K. Yadav and T. R. Singh, “Novel structural and functional impact of damaging single nucleotide polymorphisms (SNPs) on human SMYD2 protein using computational approaches,” *Meta Gene*, vol. 28, p. 100871, Jun. 2021, doi: 10.1016/j.mgene.2021.100871.

[65] S. Saxena, B. SaiAchyuth, T. Murthy, V. Chandramohan, A. Yadav, and T. Singh, “Structural and functional analysis of disease-associated mutations in GOT1 gene: An in silico study,” *ComputBiol Med.*, 2021, doi: 10.1016/j.compbiomed.2021.104695.

[66] A. Seal, A. Gupta, M. Mahalaxmi, R. Aykkal, T. R. Singh, and V. Arunachalam, “Tools, resources and databases for SNPs and indels in sequences: a review,” *Int. J. Bioinforma. Res. Appl.*, vol. 10, no. 3, pp. 264–296, Jan. 2014, doi: 10.1504/ijbra.2014.060762.

[67] “Novel structural and functional impact of damaging single nucleotide polymorphisms (SNPs) on human SMYD2 protein using computational approaches - ScienceDirect.” https://www.sciencedirect.com/science/article/pii/S2214540021000220 (accessed May 26, 2022).

[68] A. Amadei, A. B. Linssen, and H. J. Berendsen, “Essential dynamics of proteins,” *Proteins*, vol. 17, no. 4, pp. 412–425, Dec. 1993, doi: 10.1002/prot.340170408.

[69] C. C. David and D. J. Jacobs, “Principal Component Analysis: A Method for Determining the Essential Dynamics of Proteins,” *Methods Mol. Biol. Clifton NJ*, vol. 1084, pp. 193–226, 2014, doi: 10.1007/978-1-62703-658-0_11.

[70] W. Humphrey, A. Dalke, and K. Schulten, “VMD: visual molecular dynamics,” *J. Mol. Graph.*, vol. 14, no. 1, pp. 33–38, 27–28, Feb. 1996, doi: 10.1016/0263-7855(96)00018-5.

[71] E. F. Pettersen *et al.*, “UCSF Chimera--a visualization system for exploratory research and analysis,” *J. Comput. Chem.*, vol. 25, no. 13, pp. 1605–1612, Oct. 2004, doi: 10.1002/jcc.20084.