rebioin journal
OF BIOSCIENCES

IN SILICO ANALYSIS OF DELETERIOUS SNPS IN ALZHEIMER'S DISEASE GENES FOR EARLY DIAGNOSIS AND GENETIC RISK PREDICTION

admin@reboin.com

www.rebioin.com

In Silico Analysis of Deleterious SNPs in Alzheimer's Disease Genes for Early Diagnosis and Genetic Risk Prediction

Likitha AM and Annie Jessica Toppo

JSS Academy Of Higher Education And Research,Mysore, Rapture Biotech Bengaluru

Corresponding Author Email : rapturetrainer.bengaluru@gmail.com

Abstract

Alzheimer's disease (AD) is a progressive neurodegenerative disorder and the leading cause of dementia worldwide. It is characterized by cognitive impairment, memory loss, and neuropathological hallmarks such as amyloid-beta plaques and neurofibrillary tangles. Genetic factors are strongly implicated in AD pathogenesis, with several key genes—including *APOE*, *APP*, *PSEN1*, *PSEN2*, and *TREM2*—harboring single nucleotide polymorphisms (SNPs) that significantly influence disease risk. Deleterious SNPs can disrupt protein function and accelerate disease onset, making them valuable targets for early diagnostic and predictive strategies.

Traditional laboratory-based studies to evaluate SNP pathogenicity are often costly and time-intensive. In contrast, in silico approaches offer high-throughput, cost-effective alternatives for variant analysis. Computational tools such as SIFT, PolyPhen-2, PROVEAN, and MutPred2 integrate structural modeling, evolutionary conservation, and machine learning predictions to assess the functional impact of SNPs . This review discusses AD-associated genes and their deleterious variants, the classification and reported effects of SNPs, and computational pipelines for their analysis. Applications in early diagnosis, genetic risk prediction, and personalized medicine are highlighted, along with current research gaps and future directions. In silico SNP analysis holds promise as a vital tool for advancing genomic medicine and reducing the global burden of Alzheimer's disease.

Keywords: Alzheimer's disease; SNPs; deleterious variants; genetic risk prediction; bioinformatics

1. Introduction

Alzheimer's complaint represents a major health challenge, affecting over 50 million people worldwide, with prevalence projected to rise dramatically in the coming decades [1]. Clinically, announcement manifests as progressive memory decline, disabled superintendent functions, and eventual loss of independence, oppressively impacting cases and caregivers. Pathologically, announcement is characterized by extracellular amyloid- beta(Aβ) pillars, intracellular neurofibrillary befuddlements composed of hyperphosphorylated tau, and wide neuronal loss [2]. Genetics is a abecedarian factor in the etiology of Alzheimer's complaint. Domestic forms of announcement, which regard for a nonage of cases, are caused by autosomal dominant mutations in the Genes APP, PSEN1, and PSEN2 [3-4]. In discrepancy, sporadic announcement, which constitutes the maturity of cases, is explosively told by inheritable threat factors similar as the APOE ε4 allele [5-6]. Genome-wide association studies (GWAS) have linked fresh loci including CLU, PICALM, BIN1, and TREM2 — that modulate pathways related to synaptic function, lipid

metabolism, and vulnerable regulation [7-8]. Among inheritable variations, SNPs are the most abundant and extensively studied. While utmost SNPs are benign, injurious variants can profoundly affect protein function or nonsupervisory mechanisms, thereby contributing to complaint onset and progression [9]. relating similar SNPs in announcement- associated genes offers a path to uncovering molecular mechanisms and perfecting early individual and prophetic strategies. still, experimental characterization of every variant is impracticable. In this environment, in silico approaches give rapid-fire, cost-effective styles for screening SNPs and prioritizing campaigners for functional confirmation [10-12].

This study focuses on the **computational prediction of deleterious SNPs in Alzheimer's disease–associated genes**. By prioritizing variants that may disrupt protein function and regulatory pathways, we aim to bridge genomic data with disease mechanisms. Such in silico approaches provide a foundation for early diagnosis and therapeutic targeting.

2. Alzheimer's Disease Genes and Associated SNPs

2.1. Key Genes Linked to AD

Alzheimer's complaint is characterized by complex inheritable armature involving several high- impact genes that modulate complaint onset and progression. Among these, APOE is the strongest inheritable determinant of late- onset announcement. The ε4 allele of APOE increases amyloid- beta deposit, alters lipid metabolism, and confers a cure-dependent threat for developing announcement [5-6]. Domestic forms of announcement are primarily caused by autosomal dominant mutations in APP, PSEN1, and PSEN2, which affect amyloid precursor protein processing and γ- secretase exertion, leading to early- onset complaint [3-4]. Variants in TREM2, linked through genome-wide association studies, disrupt microglial function and contribute to neuroinflammation, thereby adding the threat of sporadic announcement [8]. fresh loci similar as CLU, BIN1, and PICALM have been linked to processes including synaptic function, endocytosis, and lipid homeostasis, further expanding the inheritable geography of announcement vulnerability [7].

2.2. Classification of SNPs

Single nucleotide polymorphisms(SNPs) are classified grounded on genomic position and functional impact. Synonymous SNPs do within rendering regions but don't alter the amino acid sequence. These variants can impact gene expression by affecting mRNA stability, splicing effectiveness, or restatement kinetics [12]. Non-synonymous SNPs(nsSNPs) alter the decrypted amino acid and can directly impact protein folding, stability, or function. Missense variants affect in a single amino acid negotiation, while gibberish mutations introduce a unseasonable stop codon, producing abbreviated,non-functional proteins [9]. Regulatory SNPs, located in promoters, enhancers, untranslated regions, and splice spots, can modulate recap factor list, RNA stability, or indispensable splicing patterns, thereby impacting gene expression without altering the protein sequence [14]. Intronic variants can affect splicing or recap factor reclamation, performing in altered paraphrase diversity. Both coding and nonsupervisory SNPs are applicable in announcement, as they may modify protein function(e.g., APP processing) or gene regulation(e.g., APOE expression), impacting complaint threat and progression [7-8].

2.3. Reported Deleterious SNPs in AD

Several deleterious SNPs have been experimentally characterized in AD-associated genes. The *APOE* ε4 allele (rs429358) is associated with increased disease risk and earlier onset of symptoms [5]. Mutations in *APP*, such as the V717I variant, increase amyloid-beta production and aggregation [3]. Pathogenic mutations in *PSEN1*, including M146L and E280A, result in aggressive early-onset AD with high penetrance [4]. Rare variants in *TREM2*, notably R47H, confer a three-fold increase in AD risk by impairing microglial clearance of amyloid-beta [8]. Collectively, these examples underscore the functional importance of deleterious SNPs in AD pathophysiology and highlight the need for systematic identification and characterization.

3. Computational Tools for SNP Analysis

The functional impact of SNPs can be prognosticated using bioinformatics tools. SIFT evaluates evolutionary conservation to determine whether amino acid negotiations are permitted [9]. PolyPhen- 2 assesses structural and sequence- grounded features to classify variants as benign or dangerous [10]. PROVEAN estimates the liability of injurious goods on protein function [11], while MutPred2 integrates machine literacy models to prognosticate molecular mechanisms affected by nsSNPs [12]. fresh coffers similar as CADD scores and Ensembl's Variant Effect Predictor(VEP) give comprehensive reflection [15], and databases like dbSNP and ClinVar offer expansive depositories of SNP data [16].

Below figure shows a broader view of AD genetic data analysis pipeline including preprocessing, annotation, and interpretation (Figure 1) [19]

.

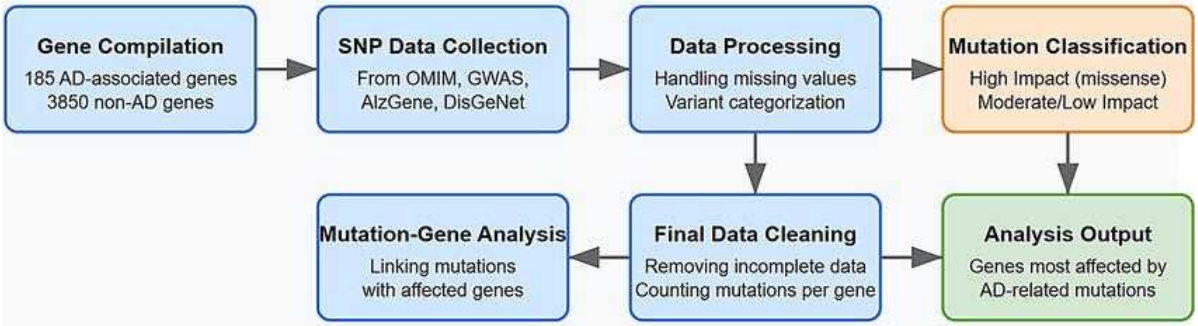

Figure 1. Six-step pipeline for Alzheimer's disease genetic data analysis.

4. Pipeline for In Silico Analysis of Deleterious SNPs

The computational analysis of deleterious SNPs follows a structured pipeline designed to prioritize variants most likely to impact protein function and disease risk. SNPs are initially retrieved from public databases such as dbSNP, ClinVar, and Ensembl. Candidate variants are filtered based on allele frequency, gene relevance, and disease association [16-15]. Functional prediction is then performed using SIFT, PolyPhen-2, PROVEAN, and MutPred2, which integrate sequence conservation, structural features, and predicted molecular mechanisms [9-12].

Structural modeling with SWISS-MODEL or I-TASSER provides insights into changes in protein conformation and stability due to SNPs, and molecular dynamics simulations assess the effects of mutations on protein folding and interactions [17]. Finally, pathway and network analyses using STRING and Cytoscape contextualize variants within biological processes and interacting proteins [18]. This integrated approach yields a prioritized set of deleterious SNPs to guide experimental validation, biomarker discovery, and therapeutic development [19].

.

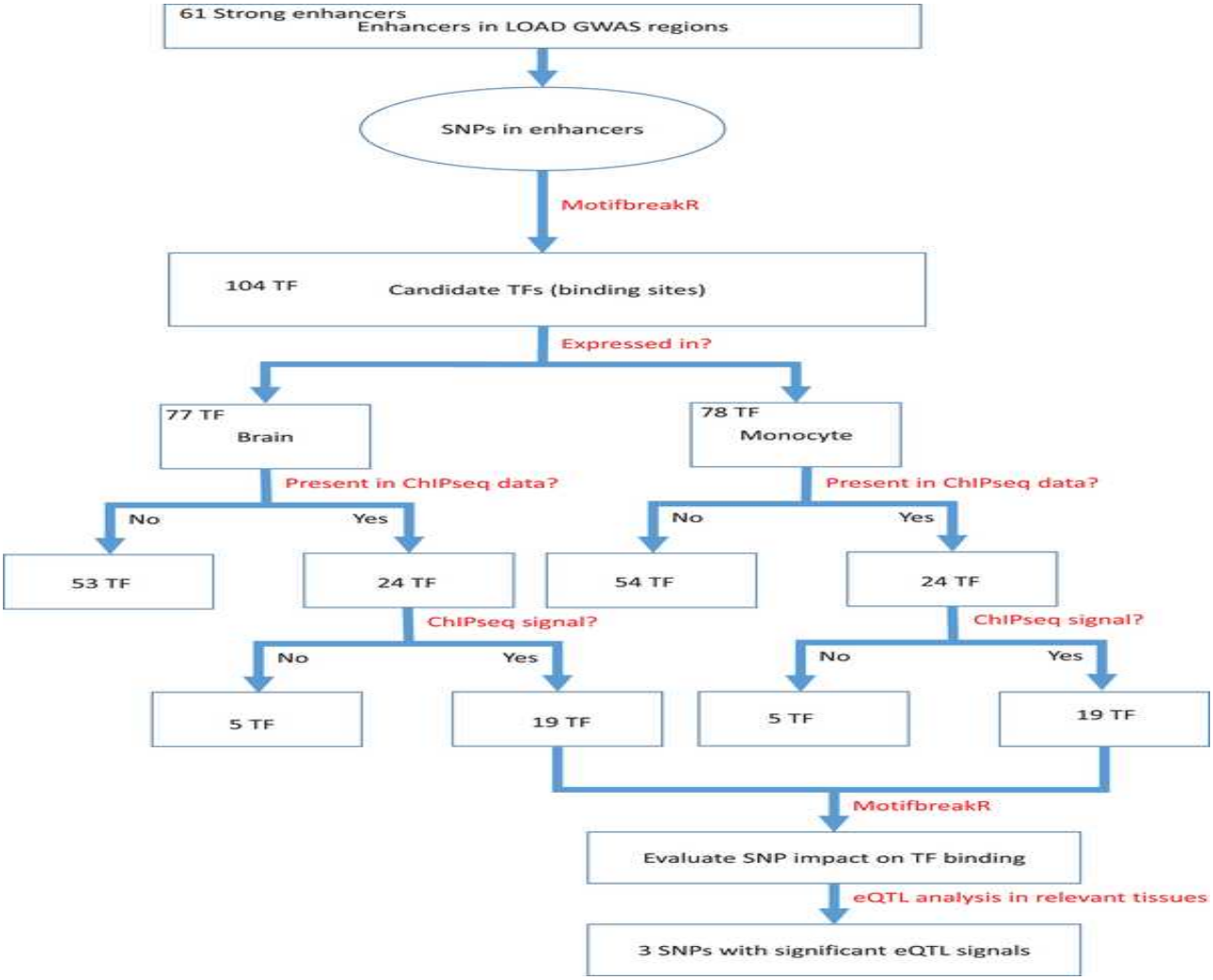

Figure 2. Bioinformatics pipeline for in silico analysis of deleterious SNPs in Alzheimer's disease-associated genes.

5. Applications in Early Diagnosis and Risk Prediction

In silico SNP analysis provides a important frame for rephrasing inheritable discoveries into clinical operations. Early opinion and threat vaticination are particularly important in announcement, as neuropathological changes begin decades before clinical symptoms [1]. Detecting injurious variants allows threat position, monitoring, and preventative interventions at preclinical stages. Prophetic inheritable testing enables rapid-fire webbing of high-threat genes similar as APOE, APP, PSEN1, and PSEN2, relating individualities at advanced inheritable threat for comforting and monitoring [5, 3, 4]. Polygenic threat scores(PRS), which combine the goods of multiple variants, enhance prophetic delicacy beyond single- gene testing and can classify individualities into low, moderate, or high-threat orders [7]. SNP analysis also aids biomarker discovery, guiding the identification of molecular autographs measurable in cerebrospinal fluid, blood, or imaging studies. Variants affecting amyloid- processing genes or

microglial function can indicate specific pathological changes for early discovery [8]. likewise, structural modeling of injurious variants facilitates the identification of remedial targets and substantiated drug approaches [12, 17].

6. Current Gaps

Despite advances, several limitations persist. Prediction discordance among tools remains common, making single-tool results unreliable [9-10]. Experimental validation of prioritized variants is resource-intensive, leading to many variants remaining as "variants of uncertain significance" (VUS) [16]. Population representation is uneven, with European cohorts overrepresented, limiting the applicability of predictive models globally [7]. Non-coding and regulatory variants are underprioritized despite their importance. Biological context, such as cell type, age, sex, and environmental factors, is insufficiently modeled. Finally, standards and clinical integration remain limited, impeding translation from computational prediction to actionable biomarkers and therapies [15, 12].

7. Future Directions

Future research aims to improve accuracy, integration, and clinical applicability. Multi-omics integration, combining genomics, transcriptomics, proteomics, metabolomics, and epigenomics, will capture downstream effects of variants and link them to clinical phenotypes [19]. Advanced AI and deep learning models are expected to refine pathogenicity predictions and identify subtle high-risk variants [12]. Expanding genomic databases to include underrepresented populations will improve global applicability [7]. Translational applications will focus on guiding experimental validation, biomarker discovery, and therapeutic development, moving in silico SNP analysis from research toward clinical genomics.

8. Conclusion

Alzheimer's disease is a complex disorder with a substantial genetic component. Deleterious SNPs in genes such as *APOE*, *APP*, *PSEN1*, *PSEN2*, and *TREM2* significantly modulate disease risk and progression. In silico approaches provide efficient and scalable methods to identify and prioritize these variants, offering insights into disease mechanisms and supporting early diagnosis and personalized medicine. Refinement of computational tools, integration of diverse datasets, and translational applications will be critical for realizing the full potential of in silico SNP analysis in AD research and clinical practice.

Acknowledgement

We would like to express our sincere gratitude to **Rapture Biotech, Bengaluru**, and **JSS College of Arts, Commerce and Science, Mysuru,** for their encouragement, guidance, and support, and for providing us with the foundation to successfully complete this work.

References

1. Alzheimer's Association. Alzheimer's Disease Facts and Figures. *Alzheimers Dement.* 2023;19(4):1598–1695. doi: 10.1002/alz.13016.

2. Selkoe DJ, Hardy J. The amyloid hypothesis of Alzheimer's disease at 25 years. *EMBO Mol Med.* 2016; 8(6):595–608. doi: 10.15252/emmm.201606210.

3. Goate A, Chartier-Harlin MC, Mullan M, Brown J, Crawford F, Fidani L, Giuffra L, Haynes A, Irving N, James L, Mant R. Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. *Nature.* 1991; 349(6311):704-6.

4. Sherrington R, Rogaev EI, Liang YA, Rogaeva EA, Levesque G, Ikeda M, Chi H, Lin C, Li G, Holman K, Tsuda T. Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. *Nature*. 1995; 375(6534):754-60.

5. Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small G, Roses AD, Haines JL, Pericak-Vance MA. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science.* 1993; 261(5123):921-3.

6. Mahley RW, Huang Y. Apolipoprotein E: from atherosclerosis to Alzheimer's disease and beyond. *Curr. Opin. Lipidol.* 1999; 10(3):207-18.

7. Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, Jun G, DeStefano AL, Bis JC, Beecham GW, Grenier-Boley B. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nature Genet.* 2013; 45(12):1452-8.

8. Guerreiro R, Wojtas A, Bras J, Carrasquillo M, Rogaeva E, Majounie E, Cruchaga C, Sassi C, Kauwe JS, Younkin S, Hazrati L. TREM2 variants in Alzheimer's disease. *N Engl J Med.* 2013; 368(2):117-27.

9. Ng PC, Henikoff S. SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res.* 2003 ;31(13):3812–4.

10. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. A method and server for predicting damaging missense mutations. *Nat. Methods.* 2010; 7(4):248-9.

11. Choi Y, Chan AP. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics.* 2015; 31(16):2745–7.

12. Pejaver V, Urresti J, Lugo-Martinez J, Pagel KA, Lin GN, Nam HJ, Mort M, Cooper DN, Sebat J, Iakoucheva LM, Mooney SD. Inferring the molecular and phenotypic impact of amino acid variants with MutPred2. *Nat. Commun.* 2020; 11(1):5918.

13. Hughes AL. Synonymous mutations: their potential impact on gene evolution. *Ann NY Acad Sci.* 2008; 1133:162–79.

14. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res.* 2012; 40(Database issue):D930–D934.

15. Cunningham F, Allen JE, Allen J, Alvarez-Jarreta J, Amode MR, Armean IM, Austine-Orimoloye O, Azov AG, Barnes I, Bennett R, Berry A. Ensembl 2022. Nucleic acids research. 2022; 50(D1):D988-95.

16. Landrum MJ, Lee JM, Benson M, Brown G, Chao C, Chitipiralla S, Gu B, Hart J, Hoffman D, Hoover J, Jang W. ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res.* 2016; 44(D1):D862-8.

17. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER Suite: protein structure and function prediction. *Nature Methods.* 2015;12(1):7-8.

18. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, Jensen LJ. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 2019;47(D1):D607-13.

19. Lutz MW, Chiba-Falek O. Bioinformatics pipeline to guide late-onset Alzheimer's disease (LOAD) post-GWAS studies: Prioritizing transcription regulatory variants within LOAD-associated regions. *Alzheimers Dement.* 2022; 8(1):e12244.

rebioin journal
OF BIOSCIENCES

IN SILICO ANALYSIS OF DELETERIOUS SNPS IN ALZHEIMER'S DISEASE GENES: ADVANCING EARLY DIAGNOSIS AND ENABLING GENETIC RISK PREDICTION FOR PREVENTIVE HEALTHCARE."

Plot no 977, GMS Road, near Balliwala Flyover, opposite Cubic Plaza,
Dehradun, Uttarakhand 248001

admin@reboin.com

www.rebioin.com