

Article

Rare Variant Analysis and Molecular Dynamics Simulation in Alzheimer's Disease Identifies Exonic Variants in FLG

Weixue Xiong ^{1,†} , Jiahui Cai ^{1,†}, Ruijia Li ², Canhong Wen ^{2,*}, Haizhu Tan ^{1,*} and on behalf of the Alzheimer's Disease Neuroimaging Initiative (ADNI) Database [‡]

¹ Department of Preventive Medicine, Shantou University Medical College, Shantou 515000, China; 20wxxiong@stu.edu.cn (W.X.); 19jhcai@stu.edu.cn (J.C.)

² Department of Statistics and Finance, School of Management, University of Science and Technology of China, Hefei 230000, China; euruijia@outlook.com

* Correspondence: wench@ustc.edu.cn (C.W.); hztan@stu.edu.cn (H.T.)

† These authors contributed equally to this work.

‡ The ADNI was launched by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies and non-profit organizations. A complete listing of ADNI investigators can be found in the acknowledgments part.



Citation: Xiong, W.; Cai, J.; Li, R.; Wen, C.; Tan, H.; on behalf of the Alzheimer's Disease Neuroimaging Initiative (ADNI) Database. Rare Variant Analysis and Molecular Dynamics Simulation in Alzheimer's Disease Identifies Exonic Variants in FLG. *Genes* **2022**, *13*, 838. <https://doi.org/10.3390/genes13050838>

Academic Editors: Jinchun Li, Chao Chen and Mingbang Wang

Received: 22 April 2022

Accepted: 4 May 2022

Published: 7 May 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Background: Although an increasing number of common variants contributing to Alzheimer's disease (AD) are uncovered by genome-wide association studies, they can only explain less than half of the heritability of AD. Rare variant association studies (RVAS) has become an increasingly important area to explain the risk or trait variability of AD. Method: To investigate the potential rare variants that cause AD, we screened 70,209 rare variants from two cohorts of a 175 AD cohort and a 214 cognitively normal cohort from the Alzheimer's Disease Neuroimaging Initiative database. MIRARE, a novel RVAS method, was performed on 232 non-synonymous variants selected by ANNOVAR annotation. Molecular docking and molecular dynamics (MD) simulation were adopted to verify the interaction between the chosen functional variants and BACE1. Results: MIRARE analysis revealed significant associations between AD and six potential pathogenic genes, including *PREX2*, *FLG*, *DHX16*, *NID2*, *ZnF585B* and *ZnF875*. Only interactions between *FLG* (including wild type and rs3120654(SER742TYR)) and *BACE1* were verified by molecular docking and MD simulation. The interaction of *FLG*(SER742TYR) with *BACE1* was greater than that of wildtype *FLG* with *BACE1*. Conclusions: According to the literature search, bio-informatics analysis, and molecular docking and MD simulation, we find non-synonymous rare variants in six genes, especially *FLG*(rs3120654), that may play key roles in AD.

Keywords: Alzheimer's disease; rare variants; ADNI; MIRARE

1. Introduction

Alzheimer's disease (AD) is a devastating illness that always causes progressive neurodegenerative and cognitive decline [1]. According to the report from WHO (key facts of dementia. Available online: <https://www.who.int/news-room/fact-sheets/detail/dementia> (accessed on 2 September 2021)) [2], there were approximately 35 million AD patients worldwide in 2021. The heritability of AD is high (79%) and more than 50% of the phenotypic variance can be explained by genetic variations [3,4]. An increasing number of common variants (minor allele frequency (MAF) > 0.05) that cause AD has been uncovered by large-scale genome-wide association studies. For example, *RBFOX1* with a higher amyloid- β burden has been linked to early AD [5]. Apolipoprotein E-4 (*APOE4*), an important possible AD therapeutic target, is highly expressed in more than half of AD patients [6]. However, these common variants explain less than half of the heritability to date [7,8]. For example, the total proportion of heritability of

APOE4 is estimated to be 25% [9,10]. With the development of base-by-base resolution of whole-genome sequencing, low-frequency variants ($0.01 < \text{MAF} < 0.05$) and rare variants ($0.005 < \text{MAF} < 0.01$), previously difficult to detect by microarrays, are becoming attractive alternative biomarkers to AD [11–13]. The reason might be that these two types of variants can explain additional disease or trait variability because the growth of the effect sizes of associated variants along with the drop of allele frequencies can exceed the limits imposed by natural selection on more common variants [11]. For example, rare variants in *TREM2* have been reported to be associated with a significant increase in the risk of AD [14], and rare variants in three genes (*APP*, *PSEN1* & *PSEN2*) can explain 5–10% of the occurrence of early onset AD [15,16]. Therefore, the identification of rare variants associated with AD is becoming increasingly important.

Due to the importance of rare variants in sequencing studies, rare variant association studies (RVAS) have become an increasingly important area to explain additional disease risk or trait variability [17,18]. Current approaches for RVAS in complex traits involve burden tests, sequence kernel association tests (SKAT), variable threshold (VT), Mixture model based Rare variant Analysis on GENes (MIRAGE), etc. In order to determine causative gene variants, MIRAGE, a novel method for determining causative gene variants, was applied in our study because it is an effective method to analyze rare variant data from case–control studies. MIRAGE uses an empirical Bayesian approach to estimate these prior probabilities by combining information across genes [19]. It not only better accounts for the heterogeneity of variant effects of genes, but also selects putative causal variants by fully using external information [19].

After selecting the significant variants for AD by using the RVAS method, it is very important to verify their relevant functions and mechanisms for AD. Over 80% of gene substrates (proteins) rarely act alone, but in complexes [20]. Protein–protein interactions (PPIs) are a useful way to predict the functionality and mechanism of a potential protein by acting with a known protein to disease [21] because proteins involved in the same cellular processes always interact with each other [21]. In an abundance of studies, amyloid β ($A\beta$) accumulation occurs before clinical symptoms in the early progression of AD [22,23]. $A\beta$ is generated by β -site APP-cleaving enzyme 1 (BACE1) by catalyzing the amyloid precursor protein (APP), and inhibition of BACE1 activity can prevent AD development [22–25]. However, natural and synthetic BACE1 inhibitors have consistently failed in the treatment of AD for reasons such as liver toxicity and lack of improvement in cognitive decline [26]. Further exploration of the pathology of BACE1 in AD is needed. Hence, BACE1 is chosen as the interaction party by several while investigating PPIs between BACE1 and the potential proteins of AD by using molecular docking, molecular dynamics (MD) simulation, etc.

The aim of our study is to explore the genetic architecture of AD. Genetic data in our study are obtained from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database (<http://adni.loni.usc.edu>, accessed on 9 September 2020) [27]. To efficiently boost the power of the RVAS and genomic evaluation accuracy, whole-genotype imputation on those preprocessed genomes was first implemented on the Michigan imputation server (<https://imputationserver.sph.umich.edu>, accessed on 25 November 2020). Then, imputed variants were functionally annotated and classified as downstream, exonic, intergenic, intronic, ncRNA-exonic, ncRNA-intronic, ncRNA splicing, splicing, and upstream by using ANNOVAR (modified on 7 October 2020) [28], using several databases, including RefGene, exac03, avsn147, dbnsfp30a, and ljb26_all. The MIRAGE method was subsequently adopted to perform RVAS analysis [19]. After that, several bio-informatics tools further provided a comprehensive demonstration of the significant selected variants. In addition, homology modeling, molecular docking, and MD simulation were performed to find the PPIs between the selected genes and BACE1 [29], and the effects of mutations on them.

2. Materials and Methods

Detailed RVAS workflow is provided in Figure 1.

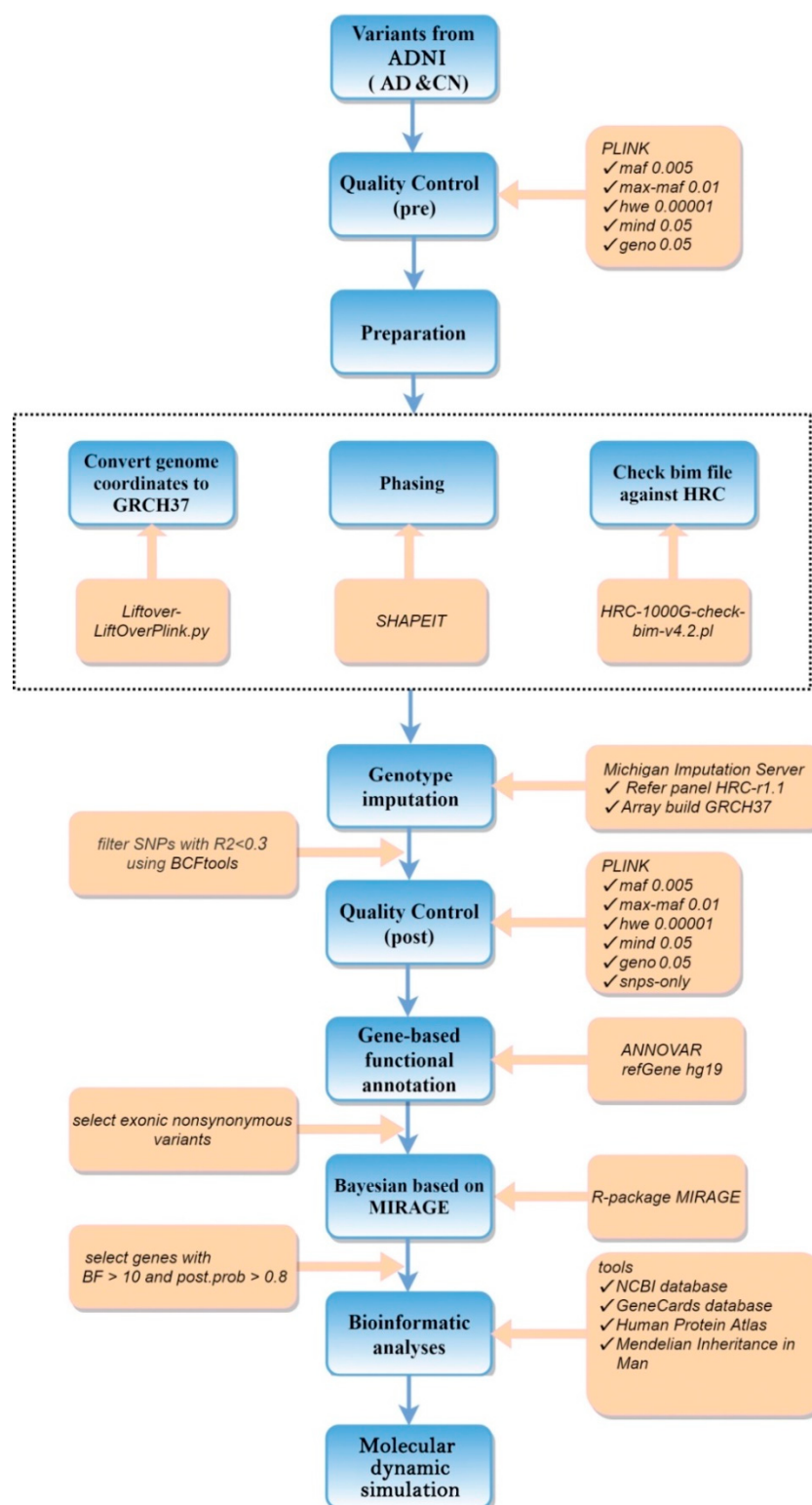


Figure 1. Rare variant association studies workflow.

2.1. Participants

Data in this study was downloaded from the ADNI-1 database, including 600,470 genetic data, several demographic and clinical data (such as age, gender, and race). Overall,

175 AD and 214 cognitively normal (CN) subjects, aged from 56.4 to 90.9 years, were enrolled. Written informed consent was obtained from all participants. The full list of inclusion and exclusion criteria can be accessed online at <https://adni.loni.usc.edu/wp-content/uploads/2008/07/adni2-procedures-manual.pdf>, accessed on 3 May 2022.

2.2. Statistical Analysis

2.2.1. Demographic Comparisons

The two-tailed student's *t*-test for two cohort comparisons in age was performed. The chi-squared test was adopted to determine the differences in gender and race distribution between two groups.

2.2.2. Data Preprocessing

A series of steps, including quality control (QC) before genotype imputation like filtering for single nucleotide polymorphism (SNPs) with call rate (missing > 5%), MAF > 1% and MAF < 0.5%), and the Hardy–Weinberg equilibrium (p -value < 1×10^{-5}), LiftOver for converting the genome coordinates into hg19, SHAPEIT2 for phasing and whole-genotype imputation on the Michigan imputation server (<https://imputationserver.sph.umich.edu>), were carried out. To filter imputed data with high degrees of uncertainty, an R-square of 0.3 was set for inclusion after imputation. Then, post-imputation QC and gene-based annotation were implemented. In this study, non-synonymous exonic variants occurring in 330 subjects were collected from the annotated variants because non-synonymous mutations were demonstrated to have a much greater effect on an individual than synonymous mutations [30]. Non-synonymous variants were subsequently predicted to be at risk or non-risk variants by considering three functionally damaging effect scores, such as polyPhen2 HDIV scores [31], Sorting Intolerant From Tolerant (SIFT) scores [32] and Combined Annotation-Dependent Depletion (CADD) scores [33] together.

2.2.3. RVAS Analysis

MIRAGE was conducted to do RVAS analysis on whole exome sequencing data of 330 subjects. Variant groups are defined depending on their annotations. Those variants with 64 PolyPhen scores lower than 0.1, a CADD score 10 or SIFT score < 0.05 were labeled as group1; others were labeled group 2. The number of the case and control groups was equal to the total number of alleles in cases and controls at a genomic locus. γ was set to 2. With the estimated parameters, the Bayes factor (BF) and posterior probabilities of each gene were computed to assess its association with the disease by comparing the likelihood of the full model (group 1) to the likelihood of the null model (group 2). In this study, a threshold of BF > 3 was set to select rare potential risk variants associated with AD.

2.3. Bioinformatic Analyses

Several in silico prediction tools were used to verify the underlying biological interpretations of the identified SNPs based on bio-informatics databases and resources, including NCBI (<https://www.ncbi.nlm.nih.gov/gene/>, accessed on 3 May 2022), GeneCards (<https://www.genecards.org/>, accessed on 3 May 2022), UCSC Genome Browser on Human (<https://genome.ucsc.edu/>, accessed on 3 May 2022), and Online Mendelian Inheritance in Man (<https://omim.org/entry/>, accessed on 3 May 2022). Online tools for the effect of SNPs on protein structure and function were also used, such as: polyPhen2 (<http://genetics.bwh.harvard.edu/pph2/index.shtml>, accessed on 3 May 2022), SIFT (<http://provean.jcvi.org/index.php>, accessed on 3 May 2022), and CADD (<https://cadd.gs.washington.edu/snv>, accessed on 3 May 2022).

2.4. Molecular Dynamics Simulation

To demonstrate whether the proteins transcribed by potential pathogenic genes were involved in the pathogenic mechanism of AD or not, MD simulations were used to verify the stability of the complex docked with BACE1, which is involved in the formation of A β . All

of the proteins that were finally chosen by MIRAGE were modeled on the trRosetta server, and the point amino acid was mutated by Chimera 1.14 [34–36]. The quality of protein structure was evaluated by PROCHECK [37]. Each selected protein was separately docked with BACE1 by the High Ambiguity Driven protein-protein Docking (HADDOCK) [38]. Finally, MD simulation was performed to test whether the above genes were involved in the formation of A β or not, and the influence of non-synonymous mutations on them.

3. Results

3.1. Demographic Comparisons

Table 1 describes the results of the difference tests between the AD and CN groups. From Table 1, we find that there were no significant differences in age (p -value = 0.907), sex (p -value = 0.815) and race (p -value = 0.357) between the two groups.

Table 1. Comparisons of demographics and clinical variables between AD and CN cohorts.

		AD	CN	p -Value
No. of participants, n (%)		175 (45%)	214 (55%)	
Sex, n (%)	M	93 (53%)	115 (54%)	0.907
	F	82 (47%)	99 (46%)	
Age, Median (IQR)		75.8 (70.85, 81.10)	75.5 (72.12, 78.38)	0.815
Race, n (%)	Asian	2 (1)	2 (1)	0.357
	Black	8 (5)	15 (7)	
	More than one	2 (1)	0 (0)	
	White	163 (93)	197 (92)	
Education, Median (IQR)		15 (12, 16)	16 (14, 18)	<0.001
Participant residence, n (%)	House	121 (69)	162 (76)	0.005
	Condo/Co-op (owned)	17 (10)	32 (15)	
	Apartment (rented)	21 (12)	13 (6)	
	Mobile Home	3 (2)	2 (1)	
	Retirement Community	5 (3)	3 (1)	
	Assisted Living	7 (4)	0 (0)	
APOE4, n (%)	Other	1 (1)	2 (1)	<0.001
	0	58 (33)	156 (73)	
	1	85 (49)	53 (25)	
	2	32 (18)	5 (2)	

3.2. RVAS Analysis

After whole-genotype imputation and QC steps, a total of 330 participants with 70,209 variants passing were reserved.

The predicted model using a data set that includes 232 non-synonymous variants in 150 ADs and 180 NCs was built by MIRAGE. Table 2 shows that two non-synonymous variants (*FLG* (BF = 4040.782) and *ZNF585B* (BF = 87.062)) are likely to be pathogenic when considering BF and post-probability together.

Table 2. The results of RAVIS analysis by using MIRAGE.

Gene	BF	Post.Prob
<i>FLG</i>	4040.782	0.996
<i>ZNF585B</i>	87.062	0.842
<i>ZNF875</i>	9.331	0.364
<i>PREX2</i>	6.746	0.293
<i>NID2</i>	5.819	0.263
<i>DHX16</i>	4.515	0.217

3.3. Molecular Dynamics Simulation

The interaction strength of the protein complex that binds is denoted by the overall HADDOCK score, which is composed of the weighted average of the Van der Waals interactions, electrostatic energy, desolvation energy, and the buried surface area (BSA). The lower the HADDOCK score is, the stronger the interaction will be. As Table 3 shows, the interaction between BACE1- mutant FLG (rs3120654 (Ser742Tyr)) is stronger than the interaction between BACE1-FLG because the HADDOCK score of BACE1-FLG (Ser742Tyr) is lower than that of BACE1-FLG. Figure 2 shows the 3D protein structure of the complex BACE1-FLG (Ser742Tyr). There was no interaction between ZNF585B and BACE1.

Table 3. Docking results of BACE1 interaction with selected proteins.

	BACE1-FLG	BACE1-FLG Ser742Tyr
HADDOCK score	-77.500 ± 13.900	-89.600 ± 16.000
Root Mean Square Deviation from the overall lowest-energy structure	0.500 ± 0.300	0.800 ± 0.500
Van der Waals energy	-79.400 ± 5.400	-83.800 ± 14.200
Electrostatic energy	-516.500 ± 81.300	-566.100 ± 71.800
Desolvation energy	2.400 ± 7.700	2.700 ± 3.400
Buried Surface Area	3250.100 ± 135.700	3318.500 ± 346.700
Z-Score	-1.400	-1.800

Notes: (1) BACE1-FLG refers to docking BACE1 with FLG; (2) BACE1-FLG (Ser742Tyr) refers to docking BACE1 with the FLG (Ser742Tyr) mutant; (3) HADDOCK score shows how well the protein complex binds with each other and is obtained from the weighted average of the Van der Waals interactions, electrostatic energy, desolvation energy, and the buried surface area.

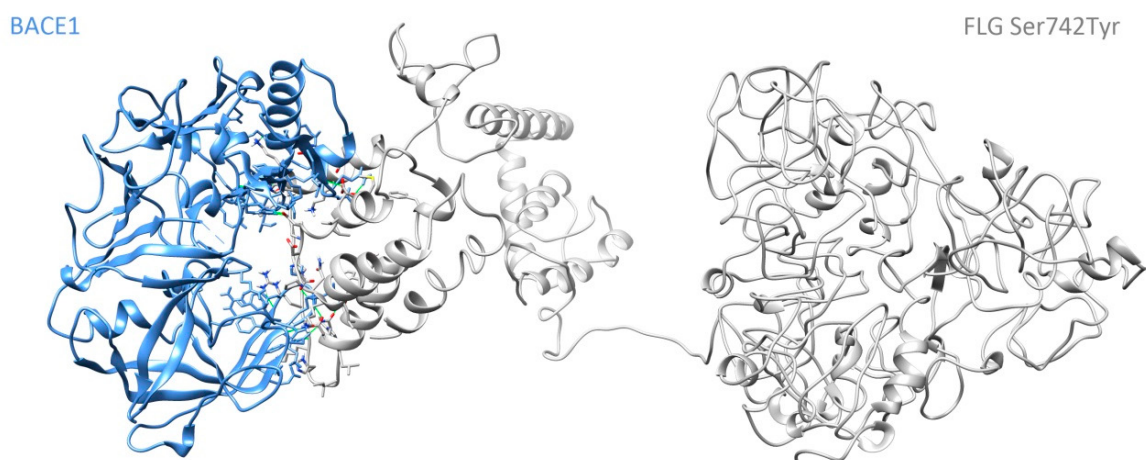


Figure 2. Three-dimensional protein structure of the complex BACE1-FLG (Ser742Tyr).

The MD simulation results are shown in Figure 3 below. The root-mean-square deviation (RMSD) plot depicts that there are no erratic fluctuations in the molecular systems and all complexes are stable (see Figure 3A). The result of the radius of gyration (Rg) in Figure 3B elucidates the volumetric and compactness variations induced in the complex. Figure 3C shows the results of solvent-accessible surface area (SASA) for the protein structures with dimensional discrepancy. From Figure 3D, we can also find that the hydrogen bonds reflect the protein's rigidity and its ability to interact with its partners (Color scheme: black: BACE1-FLG, red: BACE1-FLG (Ser742Tyr) mutant).

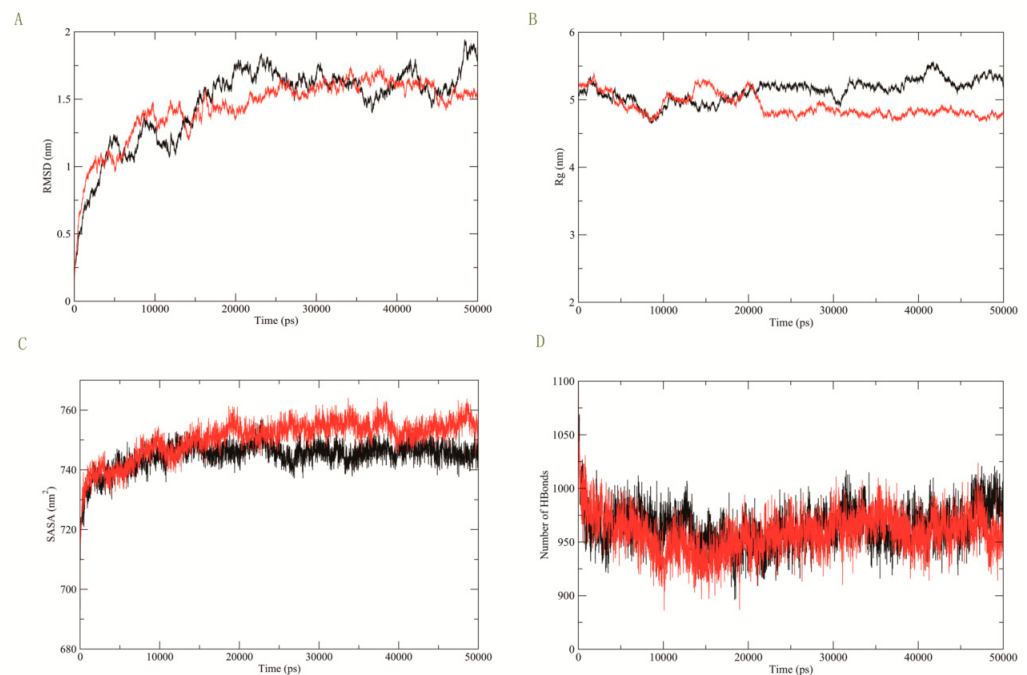


Figure 3. MD simulation results of BACE1-FLG and BACE1-FLG (Ser742Tyr) mutant. **(A)** The RMSD plot depicts that there are no erratic fluctuations in the molecular systems and all complexes are stable. **(B)** The result of the Rg elucidates the volumetric and compactness variations induced in the complex. **(C)** The result of SASA for the protein structures with dimensional discrepancy. **(D)** The hydrogen bonds reflect the protein's rigidity and its ability to interact with its partners.

4. Discussion

The results of MIRAGE showed that *PREX2*, *FLG*, *DHX16*, *NID2*, *ZnF585B*, and *ZnF875* have pathogenic risk rare variants for AD. *PREX2* is associated with brain arteriovenous malformations that can induce vascular amyloid β deposition, which is a significant risk factor for AD [39–41]. Elena Galea et al. also suggest that *PREX2* is one of the top genes detected by principal component analysis when they define AD astrocytes in multi-transcriptomic analysis [42]. The RNA helicase *Dhx16* induces transcription alterations and DNA methylation changes that are involved in memory-related neurological and neuropsychiatric diseases [43]. In single variant analysis, candidate genes including *DHX16* were identified as novel candidates for early onset AD in the study by Victoria Fernandez et al. [44]. Nidogen 2 (*NID2*) has been discovered to cause neurological disease [45]. *NID2* is a pathogenic gene for neurological disease because the severity of cerebral amyloid angiopathy is least in the striatum, where *NID2* is reduced [46]. *ZnF875* (also known as *HKR1*) is associated with aging [47]. Miren Altuna's studies find that the methylation level of *HKR1* is significantly correlated with the burden of phosphorylated tau deposits [48]. The relationship between *ZNF585B* or *ZnF875* and AD is still unknown.

Since BF greater than 10 is considered strong evidence, while a posterior probability greater than 0.8 indicates a potential efficacy signal. *FLG* (BF = 4040.782, posterior probability = 0.996) and *ZNF585B* (BF = 87.062, posterior probability = 0.842) are considered as the most likely candidate genes for rare variants [49,50]. *ZNF585B* is a P53 inhibitor, and P53 is up-regulated in AD [51,52]. *ZNF585B* is up-regulated in the olfactory bulb neural stem (OBNS) as a modulated gene, whereas OBNS enriches the ErbB signaling pathway [53]. Lack of ErbB signaling in humans has been implicated in the development of neurodegenerative diseases, such as AD [53].

In our study, filaggrin (*FLG*) was the only one of these genes with rare variants involved in the BACE1 pathway by molecular docking and MD simulation. The protein encoded by *FLG* is an intermediate filament-associated protein that aggregates keratin intermediate filaments in mammalian epidermis [54]. Shafiq M et al. suggested that

significantly higher levels of *FLG* are observed in rapidly progressive Alzheimer's disease high density fractions (HDFs) compared with sporadic Creutzfeldt-Jakob disease-specific HDFs [55]. Our bioinformatics analysis identified that three Reactome pathways for the *FLG* gene are related to developmental biology and include nervous system development (<https://reactome.org/PathwayBrowser/#/R-HSA-1266738>, accessed on 3 May 2022), formation of the cornified envelope, which is involved in axonal regeneration [56], and keratinization (Keratin 9 has been identified as an important biomarker for AD [57]).

The validation of interactions, using 50 ns (50,000 ps) MD simulations, between *FLG* and BACE1 showed stable binding between BACE1 and *FLG*. This suggests that *FLG* may affect the formation of A β by affecting the activity of BACE1, and thus participate in the pathogenesis of AD. If *FLG* was competitively bound to BACE1 or the interaction could lead to a decrease in BACE1 activity, then the competitive binding of APP to BACE1 would be decreased, leading to a reduction in A β 42 produced by hydrolysis, which can alleviate the symptoms of AD. If the interaction between *FLG* and BACE1 could increase the activity of BACE1, the activity of BACE1 would be increased, leading to an increase in A β 42 produced by hydrolysis, promoting the development of AD symptoms. Additionally, mutant *FLG* (Ser742Tyr) can affect the interaction with BACE1. These results show that the interaction between *FLG* and BACE1 is enhanced by the *FLG* (Ser742Tyr) mutation, which further supports the effect of the rare variants we found related to AD. The theoretical pathway figure is as shown in Figure 4 below.

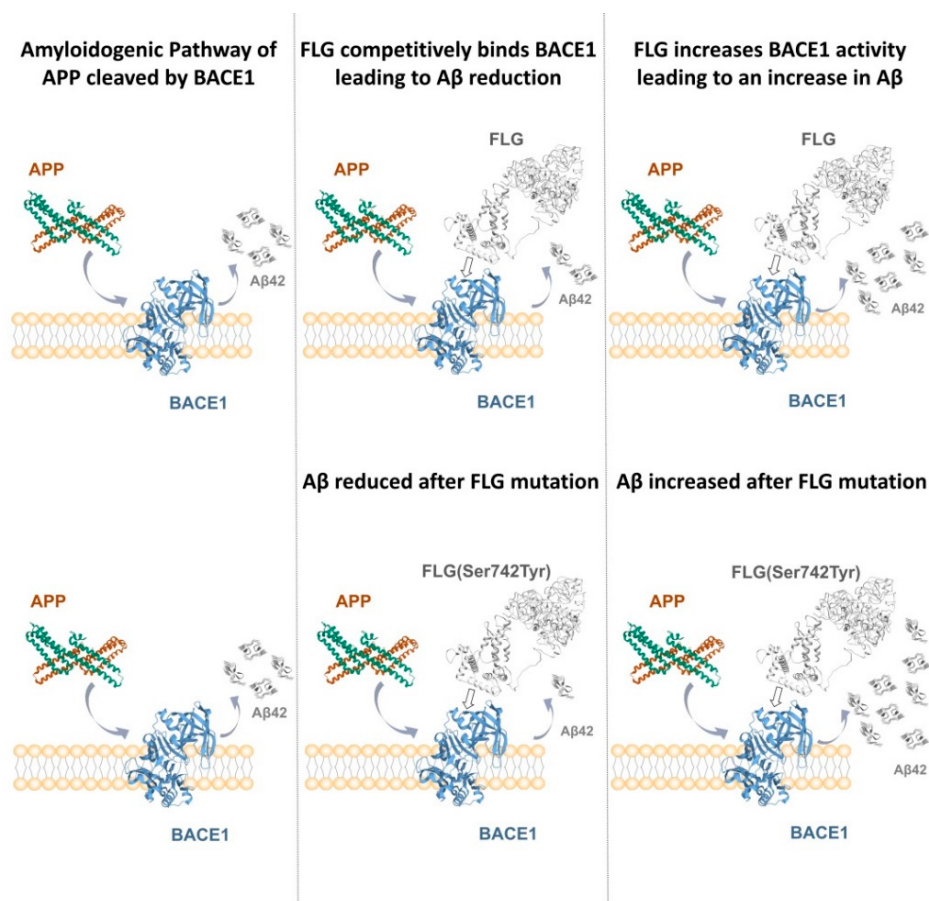


Figure 4. Hypothetical theoretical pathway figure of *FLG* and its mutant *FLG* (Ser742Tyr) on A β 42 accumulation.

Thus far, a large amount of literature, the results of bioinformatics analysis and the results of MD simulation support the existence of pathogenic rare variants in these selected genes. Only BACE1 was introduced to interact with each significant gene in our docking

and MD experiment in this study. Hence, in the future, many known proteins related with AD will be introduced to do the validation, and experiments in vitro and vivo will be performed to assess the computational work. In addition, the sample size in our study is not large enough for our RVAS research.

In conclusion, we identified six genes with rare variants affecting AD, especially *FLG*. We hope our study may aid in furthering study of exploring the underlying pathology mechanism of AD.

Author Contributions: H.T. and C.W. conceived and oversaw the project. W.X., J.C. and R.L. contributed to the data analysis and MD simulation and performed the bioinformatics analysis. All authors made critical contributions to the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: Our research was supported by the National Key Research and Development Program of China (Reference: 2018YFC1315400); the Science and Technology Planning Project of Guangdong Province (Reference: 2017A010101030); the Third Medical Technology Projects of Shantou in 2018 (Reference: 41368043); the Natural Science Foundation of China (References: 11771462, 71991474, 72171216), the National Key Research, the Key Research and Development Program of Guangdong, China (Reference: 2019B020228001), the Science and Technology Program of Guangzhou, China (Grant No. 202002030129), and the Natural Science Foundation of Anhui (Reference: BJ2040170017).

Institutional Review Board Statement: Publicly available databases were analyzed in this study. The ADNI database was conducted in accordance with the Declaration of Helsinki, and approved by the ethics committees of the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and the local ethics committee of each participating site (<https://adni.loni.usc.edu/methods/documents/>, accessed on 3 May 2022).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study. Written informed consent has been obtained from the patient(s) to publish this paper.

Data Availability Statement: Data in our study are obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (<http://adni.loni.usc.edu>, accessed on 3 May 2022), an open source database.

Acknowledgments: We gratefully thank Geng Dong for his support in computational biology and bioinformatics technology. The authors would like to thank blogger Wenyan Chen and Mo Xv for their help. Thanks to investigators within ADNI such as Michael W. Weiner who contributed to the design and implementation of ADNI and provided data, but did not participate in the analysis or writing of this report. A complete listing of ADNI investigators can be found at: https://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf, accessed on 3 May 2022.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Aggleton, J.P.; Pralus, A.; Nelson, A.J.; Hornberger, M. Thalamic pathology and memory loss in early Alzheimer's disease: Moving the focus from the medial temporal lobe to Papez circuit. *Brain* **2016**, *139*, 1877–1890. [CrossRef] [PubMed]
2. Key Facts of Dementia. Available online: <https://www.who.int/news-room/fact-sheets/detail/dementia> (accessed on 2 September 2021).
3. Gatz, M.; Reynolds, C.A.; Fratiglioni, L.; Johansson, B.; Mortimer, J.A.; Berg, S.; Fiske, A.; Pedersen, N.L. Role of genes and environments for explaining Alzheimer disease. *Arch. Gen. Psychiatry* **2006**, *63*, 168–174. [CrossRef] [PubMed]
4. Ridge, P.G.; Hoyt, K.B.; Boehme, K.; Mukherjee, S.; Crane, P.K.; Haines, J.L.; Mayeux, R.; Farrer, L.A.; Pericak-Vance, M.A.; Schellenberg, G.D. Assessment of the genetic variance of late-onset Alzheimer's disease. *Neurobiol. Aging* **2016**, *41*, 200.e213–200.e220. [CrossRef] [PubMed]
5. Raghavan, N.S.; Dumitrescu, L.; Mormino, E.; Mahoney, E.R.; Lee, A.J.; Gao, Y.; Bilgel, M.; Goldstein, D.; Harrison, T.; Engelman, C.D. Association between common variants in RBFox1, an RNA-binding protein, and brain amyloidosis in early and preclinical Alzheimer disease. *JAMA Neurol.* **2020**, *77*, 1288–1298. [CrossRef]
6. Michaelson, D.M. APOE ϵ 4: The most prevalent yet understudied risk factor for Alzheimer's disease. *Alzheimer's Dement.* **2014**, *10*, 861–868. [CrossRef]

7. Zuk, O.; Schaffner, S.F.; Samocha, K.; Do, R.; Hechter, E.; Kathiresan, S.; Daly, M.J.; Neale, B.M.; Sunyaev, S.R.; Lander, E.S. Searching for missing heritability: Designing rare variant association studies. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E455–E464. [\[CrossRef\]](#)
8. Cruchaga, C.; Chakraverty, S.; Mayo, K.; Vallania, F.L.; Mitra, R.D.; Faber, K.; Williamson, J.; Bird, T.; Diaz-Arrastia, R.; Foroud, T.M. Rare variants in APP, PSEN1 and PSEN2 increase risk for AD in late-onset Alzheimer's disease families. *PLoS ONE* **2012**, *7*, e31039. [\[CrossRef\]](#)
9. So, H.C.; Gui, A.H.; Cherny, S.S.; Sham, P.C. Evaluating the heritability explained by known susceptibility variants: A survey of ten complex diseases. *Genet. Epidemiol.* **2011**, *35*, 310–317. [\[CrossRef\]](#)
10. Bagyinszky, E.; Giau, V.V.; An, S.A. Transcriptomics in Alzheimer's Disease: Aspects and Challenges. *Int. J. Mol. Sci.* **2020**, *21*, 3517. [\[CrossRef\]](#)
11. Cirulli, E.T.; White, S.; Read, R.W.; Elhanan, G.; Metcalf, W.J.; Tanudjaja, F.; Fath, D.M.; Sandoval, E.; Isaksson, M.; Schlauch, K.A. Genome-wide rare variant analysis for thousands of phenotypes in over 70,000 exomes from two cohorts. *Nat. Commun.* **2020**, *11*, 542. [\[CrossRef\]](#)
12. Huffman, J.E.; De Vries, P.S.; Morrison, A.C.; Sabater-Lleal, M.; Kacprowski, T.; Auer, P.L.; Brody, J.A.; Chasman, D.I.; Chen, M.-H.; Guo, X. Rare and low-frequency variants and their association with plasma levels of fibrinogen, FVII, FVIII, and vWF. *Blood* **2015**, *126*, e19–e29. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Manolio, T.A.; Collins, F.S.; Cox, N.J.; Goldstein, D.B.; Hindorf, L.A.; Hunter, D.J.; McCarthy, M.I.; Ramos, E.M.; Cardon, L.R.; Chakravarti, A. Finding the missing heritability of complex diseases. *Nature* **2009**, *461*, 747–753. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Guerreiro, R.; Wojtas, A.; Bras, J.; Carrasquillo, M.; Rogaeva, E.; Majounie, E.; Cruchaga, C.; Sassi, C.; Kauwe, J.S.; Younkin, S. TREM2 variants in Alzheimer's disease. *N. Engl. J. Med.* **2013**, *368*, 117–127. [\[CrossRef\]](#)
15. Van Cauwenberghe, C.; Van Broeckhoven, C.; Sleegers, K. The genetic landscape of Alzheimer disease: Clinical implications and perspectives. *Genet. Med.* **2016**, *18*, 421–430. [\[CrossRef\]](#)
16. Hollingworth, P.; Harold, D.; Sims, R.; Gerrish, A.; Lambert, J.-C.; Carrasquillo, M.M.; Abraham, R.; Hamshere, M.L.; Pahwa, J.S.; Moskva, V. Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease. *Nat. Genet.* **2011**, *43*, 429–435. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Liu, L.; Wang, P.; Meng, J.; Chen, L.; Zhu, W.; Ma, W. A permutation method for detecting trend correlations in rare variant association studies. *Genet. Res.* **2019**, *101*, E13. [\[CrossRef\]](#)
18. Lee, S.; Abecasis, G.R.; Boehnke, M.; Lin, X. Rare-variant association analysis: Study designs and statistical tests. *Am. J. Hum. Genet.* **2014**, *95*, 5–23. [\[CrossRef\]](#)
19. Han, S.; Knoblauch, N.; Wang, G.; Zhao, S.; Liu, Y.; Xie, Y.; Sheng, W.; Nguyen, H.T.; He, X. A Bayesian method for rare variant analysis using functional annotations and its application to Autism. *bioRxiv* **2019**, 828061.
20. Hu, G.; Di Paola, L.; Liang, Z.; Giuliani, A. Comparative study of elastic network model and protein contact network for protein complexes: The hemoglobin case. *BioMed Res. Int.* **2017**, *2017*, 2483264. [\[CrossRef\]](#)
21. Von Mering, C.; Krause, R.; Snel, B.; Cornell, M.; Oliver, S.G.; Fields, S.; Bork, P. Comparative assessment of large-scale data sets of protein–protein interactions. *Nature* **2002**, *417*, 399–403. [\[CrossRef\]](#)
22. Koelsch, G. BACE1 function and inhibition: Implications of intervention in the amyloid pathway of Alzheimer's disease pathology. *Molecules* **2017**, *22*, 1723. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Yan, R.; Vassar, R. Targeting the β secretase BACE1 for Alzheimer's disease therapy. *Lancet Neurol.* **2014**, *13*, 319–329. [\[CrossRef\]](#)
24. Franzmeier, N.; Dewenter, A.; Frontzkowski, L.; Dichgans, M.; Rubinski, A.; Neitzel, J.; Smith, R.; Strandberg, O.; Ossenkoppele, R.; Burger, K. Patient-centered connectivity-based prediction of tau pathology spread in Alzheimer's disease. *Sci. Adv.* **2020**, *6*, eabd1327. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Swaminathan, G.; Zhu, W.; Plowey, E.D. BECN1/Beclin 1 sorts cell-surface APP/amyloid β precursor protein for lysosomal degradation. *Autophagy* **2016**, *12*, 2404–2419. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Chakravarthy, M.; Veedu, R.N. BACE1 inhibition using 2'-OMePS steric blocking antisense oligonucleotides. *Genes* **2019**, *10*, 705. [\[CrossRef\]](#)
27. Spooner, A.; Chen, E.; Sowmya, A.; Sachdev, P.; Kochan, N.A.; Trollor, J.; Brodaty, H. A comparison of machine learning methods for survival analysis of high-dimensional clinical data for dementia prediction. *Sci. Rep.* **2020**, *10*, 1–10. [\[CrossRef\]](#)
28. Wang, K.; Li, M.; Hakonarson, H. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **2010**, *38*, e164. [\[CrossRef\]](#)
29. Ghosh, A.K.; Osswald, H.L. BACE1 (β -secretase) inhibitors for the treatment of Alzheimer's disease. *Chem. Soc. Rev.* **2014**, *43*, 6765–6813. [\[CrossRef\]](#)
30. Agashe, D.; Sane, M.; Phalnikar, K.; Diwan, G.D.; Habibullah, A.; Martinez-Gomez, N.C.; Sahasrabudhe, V.; Polachek, W.; Wang, J.; Chubiz, L.M. Large-effect beneficial synonymous mutations mediate rapid and parallel adaptation in a bacterium. *Mol. Biol. Evol.* **2016**, *33*, 1542–1553. [\[CrossRef\]](#)
31. Adzhubei, I.A.; Schmidt, S.; Peshkin, L.; Ramensky, V.E.; Gerasimova, A.; Bork, P.; Kondrashov, A.S.; Sunyaev, S.R. A method and server for predicting damaging missense mutations. *Nat. Methods* **2010**, *7*, 248–249. [\[CrossRef\]](#)
32. Ng, P.C.; Henikoff, S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res.* **2003**, *31*, 3812–3814. [\[CrossRef\]](#) [\[PubMed\]](#)

33. Kircher, M.; Witten, D.M.; Jain, P.; O’Roak, B.J.; Cooper, G.M.; Shendure, J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.* **2014**, *46*, 310–315. [[CrossRef](#)] [[PubMed](#)]
34. Yang, J.; Yan, R.; Roy, A.; Xu, D.; Poisson, J.; Zhang, Y. The I-TASSER Suite: Protein structure and function prediction. *Nat. Methods* **2015**, *12*, 7–8. [[CrossRef](#)] [[PubMed](#)]
35. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612. [[CrossRef](#)]
36. Bassot, C.; Elofsson, A. Accurate contact-based modelling of repeat proteins predicts the structure of new repeats protein families. *PLoS Comput. Biol.* **2021**, *17*, e1008798. [[CrossRef](#)]
37. Laskowski, R.A.; MacArthur, M.W.; Moss, D.S.; Thornton, J.M. PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **1993**, *26*, 283–291. [[CrossRef](#)]
38. Dominguez, C.; Boelens, R.; Bonvin, A.M. HADDOCK: A protein–Protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.* **2003**, *125*, 1731–1737. [[CrossRef](#)]
39. Wang, K.; Zhao, S.; Liu, B.; Zhang, Q.; Li, Y.; Liu, J.; Shen, Y.; Ding, X.; Lin, J.; Wu, Y. Perturbations of BMP/TGF- β and VEGF/VEGFR signalling pathways in non-syndromic sporadic brain arteriovenous malformations (BAVM). *J. Med. Genet.* **2018**, *55*, 675–684. [[CrossRef](#)]
40. Hart, M.; Merz, P.; Bennett-Gray, J.; Menezes, A.; Goeken, J.; Schelper, R.; Wisniewski, H. β -amyloid protein of Alzheimer’s disease is found in cerebral and spinal cord vascular malformations. *Am. J. Pathol.* **1988**, *132*, 167.
41. Tsonou, E. Exploring the Importance of PREX Proteins in Glucose Homeostasis and Insulin Signalling. Ph.D. Thesis, University of Cambridge, Cambridge, UK, 2020.
42. Galea, E.; Weinstock, L.D.; Larramona-Arcas, R.; Pybus, A.F.; Giménez-Llort, L.; Escartin, C.; Wood, L.B. Multi-transcriptomic analysis points to early organelle dysfunction in human astrocytes in Alzheimer’s disease. *Neurobiol. Dis.* **2022**, *166*, 105655. [[CrossRef](#)]
43. Duke, C.G.; Kennedy, A.J.; Gavin, C.F.; Day, J.J.; Sweatt, J.D. Experience-dependent epigenomic reorganization in the hippocampus. *Learn. Mem.* **2017**, *24*, 278–288. [[CrossRef](#)]
44. Fernandez, V.; Huey, D.; Budde, J.P.; Farias, F.H.; Harari, O.; Norton, J.; Gentsch, J.; Morris, J.C.; Goate, A.; Rademakers, R. O2-10-02: Whole Exome Sequencing Analysis in Early Onset Alzheimer Reveals Novel Candidate Genes. *Alzheimer’s Dement.* **2019**, *15*, P564. [[CrossRef](#)]
45. Li, C. Understanding Rare Neurological Disorders Using Drosophila Models: Mechanistic Characterization of Neurotoxicity in Snyder-Robinson Syndrome. Ph.D. Thesis, University of Miami, Coral Gables, FL, USA, 2018.
46. Hawkes, C.A.; Gatherer, M.; Sharp, M.M.; Dorr, A.; Yuen, H.M.; Kalaria, R.; Weller, R.O.; Carare, R.O. Regional differences in the morphological and functional effects of aging on cerebral basement membranes and perivascular drainage of amyloid- β from the mouse brain. *Aging Cell* **2013**, *12*, 224–236. [[CrossRef](#)] [[PubMed](#)]
47. Zeng, Q.; Chen, X.; Ning, C.; Zhu, Q.; Yao, Y.; Zhao, Y.; Luan, F. Methylation of the genes ROD1, NLRC5, and HKR1 is associated with aging in Hainan centenarians. *BMC Med. Genom.* **2018**, *11*, 7. [[CrossRef](#)] [[PubMed](#)]
48. Altuna, M.; Urdanoz-Casado, A.; Sánchez-Ruiz de Gordo, J.; Zelaya, M.V.; Labarga, A.; Lepesant, J.M.; Roldán, M.; Blanco-Luquin, I.; Perdonés, Á.; Larumbe, R. DNA methylation signature of human hippocampus in Alzheimer’s disease is linked to neurogenesis. *Clin. Epigenetics* **2019**, *11*, 91. [[CrossRef](#)]
49. van Doorn, J.; van den Bergh, D.; Böhm, U.; Dablander, F.; Derks, K.; Draws, T.; Etz, A.; Evans, N.J.; Gronau, Q.F.; Haaf, J.M. The JASP guidelines for conducting and reporting a Bayesian analysis. *Psychon. Bull. Rev.* **2021**, *28*, 813–826. [[CrossRef](#)]
50. Hewett, K.; Sanders, D.B.; Grove, R.A.; Broderick, C.L.; Rudo, T.J.; Bassiri, A.; Zvartau-Hind, M.; Bril, V.; Group, B.S. Randomized study of adjunctive belimumab in participants with generalized myasthenia gravis. *Neurology* **2018**, *90*, e1425–e1434. [[CrossRef](#)]
51. Hooper, C.; Meimaridou, E.; Tavassoli, M.; Melino, G.; Lovestone, S.; Killick, R. p53 is upregulated in Alzheimer’s disease and induces tau phosphorylation in HEK293a cells. *Neurosci. Lett.* **2007**, *418*, 34–37. [[CrossRef](#)]
52. Black, S.K. An Investigation into the Roles of p53, Nodal/Activin and Fibroblast Growth Factor Signalling in Early Heart Development. Ph.D. Thesis, Cardiff University, Cardiff, UK, 2016.
53. Marei, H.E.; Althani, A.; Afifi, N.; Abd-Elmaksoud, A.; Bernardini, C.; Michetti, F.; Barba, M.; Pescatori, M.; Maira, G.; Paldino, E. Over-expression of hNGF in adult human olfactory bulb neural stem cells promotes cell growth and oligodendrocytic differentiation. *PLoS ONE* **2013**, *8*, e82206. [[CrossRef](#)]
54. Soltani, S.; Saghaadeh, A.; Movahedi, M.; Tavakol, M.; Sadr, M.; Farhadi, E.; Rezaei, N. FLG single nucleotide polymorphisms in chronic idiopathic urticaria. *Allergol. Immunopathol.* **2016**, *44*, 341–345. [[CrossRef](#)]
55. Shafiq, M.; Zafar, S.; Younas, N.; Noor, A.; Puig, B.; Altmepfen, H.C.; Schmitz, M.; Matschke, J.; Ferrer, I.; Glatzel, M. Prion protein oligomers cause neuronal cytoskeletal damage in rapidly progressive Alzheimer’s disease. *Mol. Neurodegener.* **2021**, *16*, 11. [[CrossRef](#)] [[PubMed](#)]
56. Bonilla, I.E.; Tanabe, K.; Strittmatter, S.M. Small proline-rich repeat protein 1A is expressed by axotomized neurons and promotes axonal outgrowth. *J. Neurosci.* **2002**, *22*, 1303–1315. [[CrossRef](#)] [[PubMed](#)]
57. Richens, J.L.; Spencer, H.L.; Butler, M.; Cantlay, F.; Vere, K.-A.; Bajaj, N.; Morgan, K.; O’Shea, P. Rationalising the role of Keratin 9 as a biomarker for Alzheimer’s disease. *Sci. Rep.* **2016**, *6*, 22962. [[CrossRef](#)] [[PubMed](#)]