DOI:10.25092/baunfbed. 1197932 *J. BAUN Inst. Sci. Technol., 25(2), 489-500, (2023)*

Evaluation of missense SNVs within human *APOE* (Apolipoprotein E) gene via bioinformatics tools

Ömer Faruk KARASAKAL[1*], Ebru ÖZKAN OKTAY[2], Tuğba KAMAN[3]

[1]*Üsküdar University, Vocational School of Health Services, Medical Laboratory Techniques, Üsküdar, Istanbul*
[2]*Üsküdar University, Vocational School of Health Services, Laboratory Technology, Üsküdar, Istanbul*
[3]*Üsküdar University, Vocational School of Health Services, Medical and Aromatic Plants, Üsküdar, Istanbul*

Geliş Tarihi (Received Date): 01.11.2022
Kabul Tarihi (Accepted Date): 05.04.2023

Abstract

Apolipoprotein E (APOE) is one of the main proteins responsible for cholesterol transport. It has three major isoforms, APOE2, APOE3, and APOE4. The purpose of this study is to investigate the possible effects of single nucleotide variations (SNVs) in the APOE gene, which cause amino acid substitution, on the function, structure and stabilization of the APOE protein using bioinformatics/s tools. SNVs and protein sequence information were obtained from NCBI and UniProt databases. Bioinformatical analysis was performed using a series of tools such as SIFT, PolyPhen-2, SNPs&GO, Mutation Assessor, PROVEAN, SNAP2, I-Mutant-3, MUPro, and Project HOPE. As a result, 321 missense SNVs were analyzed and rs7412 (R176C), rs769455 (R163C), rs11542029 (R50C), rs121918393 (R154S), rs121918394 (K164Q), rs200703101 (R154P), rs387906567 (R160C), rs11542040 (P102T), rs11542041 (R132S) and rs41382345 (E139V) were predicted to be deleterious/disease related after functional analysis and pathological effect analysis via all of the bioinformatics/s tools. According to the protein stabilization results, it was determined that all SNVs decreased protein stabilization with the MUPro software tool, and two SNVs (rs121918394, rs41382345) increased protein stabilization with the I-Mutant-3 software tool. The models of protein and amino acid properties were obtained via Project HOPE for all high-risk SNVs. We hope our analysis will be valuable for further proteomic, genomic, and clinical research.

Keywords: *Apolipoprotein E (APOE), cholesterol, in silico, single nucleotide variation (SNV)*

[*]Ömer Faruk Karasakal, omerfaruk.karasakal@uskudar.edu.tr, https://orcid.org/0000-0001-7803-3249,
Ebru Özkan Oktay, ebru.ozkanoktay@uskudar.edu.tr, https://orcid.org/0000-0002-0395-9845
Tuğba Kaman, tuğba.kaman@uskudar.edu.tr, https://orcid.org/0000-0002-5885-0193

Biyoinformatik araçlar aracılığıyla insan *APOE* (Apolipoprotein E) genindeki yanlış anlamlı SNV'lerin değerlendirilmesi

Öz

Apolipoprotein E (APOE), kolesterol taşınmasından sorumlu ana proteinlerden biridir. APOE2, APOE3 ve APOE4 olmak üzere üç ana izoformu vardır. Bu çalışmanın amacı, APOE geninde amino asit yer değiştirmesine neden olan tek nükleotid varyantlarının (SNV'ler) APOE proteininin işlevi, yapısı ve stabilizasyonu üzerindeki olası etkilerini biyoinformatik araçlar kullanarak araştırmaktır. SNV'ler ve protein dizi bilgisi, NCBI ve UniProt veritabanlarından elde edilmiştir. Biyoinformatik analiz, SIFT, PolyPhen-2, SNPs&GO, Mutation Assessor, PROVEAN, SNAP2, I-Mutant-3, MUPro ve Project HOPE gibi bir dizi araç kullanılarak yapılmıştır. Sonuç olarak 321 yanlış anlamlı SNV analiz edilmiş ve rs7412 (R176C), rs769455 (R163C), rs11542029 (R50C), rs121918393 (R154S), rs121918394 (K164Q), rs200703101 (R154P), rs387906567 (R160C), 102rsT11542040 rs11542041 (R132S) ve rs41382345 (E139V)'nin tüm biyoinformatik araçları aracılığıyla fonksiyonel analiz ve patolojik etki analizinden sonra zararlı/hastalık ile ilgili olduğu tahmin edilmiştir. Protein stabilizasyon sonuçlarına göre, MUPro yazılım aracı ile tüm SNV'lerin protein stabilizasyonunu azalttığı ve I-Mutant-3 yazılım aracı ile iki SNV'nin (rs121918394, rs41382345) protein stabilizasyonunu arttırdığı belirlenmiştir. Protein ve amino asit özellikleri modelleri, tüm yüksek riskli SNV'ler için Project HOPE aracılığıyla elde edilmiştir. Analizimizin daha fazla proteomik, genomik ve klinik araştırmalar için değerli olacağını umuyoruz.

Anahtar Kelimeler: *Apolipoprotein E (APOE), kolesterol, in silico, tek nükleotid varyantı (SNV)*

1. Introduction

Apolipoprotein E (APOE) is one of the main proteins responsible for cholesterol transport. APOE2, APOE3, and APOE4 are main isoforms of APOE [1]. These isoforms are similar except 112^{th} and 158^{th} amino acids [2]. It plays several important roles in the regulation of plasma lipid and lipoprotein levels. APOE is a component of plasma lipoproteins. One of the main function is to mediate high-affinity binding of APOE-containing lipoproteins to the low density lipoprotein (LDL) receptor and the LDL receptor-related protein (LRP) and to cell-surface heparan sulfate proteoglycans (HSPG) [3]. In the literature, it has been focused on the relationship between the severity and mortality of COVID-19 disease and the APOE ε4/ε4 genotypes of individuals. It has been reported that clinical studies are required to understand the biological mechanisms related to APOE genotypes and COVID-19 severity [4-5]. Furthermore, Alzheimer's disease, infectious diseases, atherosclerosis, type III hyperlipoproteinemia, impaired cognitive function, and telomere shortening have been associated with APOE isoforms [6-11]. The aim of this study is to evaluate a range of effects of single nucleotide *variations* (SNVs) in the APOE gene, which cause amino acid substitution, on the structure, function and stabilization of the APOE protein via bioinformatics/s tools.

2. Material and methods

Bioinformatical analysis was performed using a series of tools as described previous studies [12]; [13]. The workflow and software tools used in our study are presented in the figure 1.

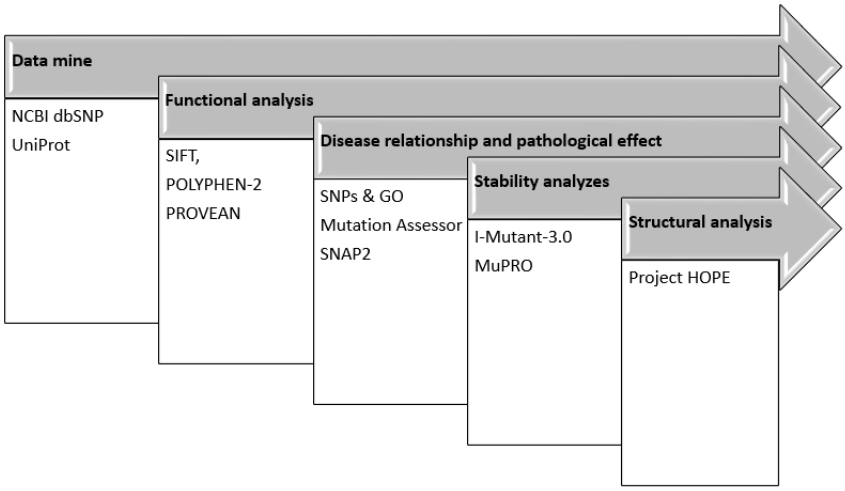


Figure 1. The schema and software tools

Missense SNVs in the *APOE* gene were retrieved from the NCBI dbSNV database. The information of the amino acid substitutions caused by those SNVs were obtained and used in further analysis (https://www.ncbi.nlm.nih.gov/SNV/). The UniProtKB number, entry name, and FASTA format protein sequence were provided from UniProt database and used as input data which were required for further analysis (https://www.uniprot.org/). SIFT (Sorting Intolerant From Tolerant), PolyPhen-2 (Polymorphism Phenotyping-2), and PROVEAN (Protein Variation Effect Analyzer) software tools were used to investigate whether amino acid changes will alter protein function or not. The deleterious/damaging effects of SNVs were predicted using SIFT (https://sift.bii.a-star.edu.sg/) [14], PolyPhen-2 (Both HumVar and HumDiv were used) (http://genetics.bwh.harvard.edu/pph2/) [15], and PROVEAN (http://provean.jcvi.org/index.php) [16] software tools. The functional impacts of amino acid alterations in proteins were estimated via Mutation Assessor (http://mutationassessor.org/r3/) [17] and SNAP2 (https://rostlab.org/services/snap2web/) servers. SNPs&GO (https://SNVs-and-go.biocomp.unibo.it/SNVs-and-go/) server was used to obtain disease-related SNVs in proteins with functional annotations [18]. The protein stability changes owing to amino acid substitutions were predicted by I-Mutant 3.0 (http://gpcr2.biocomp.unibo.it/cgi/predictors/I-Mutant3.0/I-Mutant3.0.cgi) [19] and MUpro (http://mupro.proteomics.ics.uci.edu/) software tools [20]. Finally, Project HOPE (https://www3.cmbi.umcn.nl/hope/) software tool was used to obtain 3D models of protein and polymorphism sites [21].

3. Results

SNVs in the *APOE* gene were procured from the NCBI dbSNV in December 2020. It has been determined that there are a total of 2003 SNVs in the *APOE* gene, and 321 of them are missense SNVs. Among the 321 SNVs, the SNVs that were found to be

deleterious/damaging in all programs were considered as high-risk SNVs. In silico prediction results is shown in Table 1, stabilization analysis results are shown in Table 2. Three-dimensional models are presented in Table 3. When amino acids are evaluated in terms of size, wild-type residues of R176C, R163C, R50C, R154S, K164Q, R154P, R160C, R132S, E139V are shown to be larger than mutant residues. The mutant residue is smaller than the other residue will result in a possible loss of outer interaction. No size difference is specified for position P102T [21].

When evaluated in terms of charge of amino acid residues, wild type residues have positive charge whereas mutant type residues have neutral charge in rs7412 (R176C), rs769455 (R163C), rs11542029 (R50C), rs121918393 (R154S), rs121918394 (K164Q), rs200703101 (R154P), rs387906567 (R160C), rs11541 (R132). The electrical charge at position rs41382345 (E139V) is indicated as negative for wild residue and neutral for mutant residue. As a result, this can result in loss of interaction with other molecules. However, no electrical load difference is reported for position rs11542040 (P102T) [21]. When evaluated in terms of hydrophobicity, mutant residues in rs7412 (R176C), rs769455 (R163C), rs11542029 (R50C), rs121918393 (R154S), rs200703101 (R154P), rs387906567 (R160C), rs11542041 (R132S), rs41382345 (E139V) are more hydrophobic than the wild-type residue. It is stated that the difference in hydrophobicity of these residues will disrupt the true folding. For rs11542040 (P102T), the wild residue was reported to be more hydrophobic than the other residue. The alteration can cause to loss of hydrophobic interactions with other molecules on the protein surface. No difference in hydrophobicity is noted for position rs121918394 (K164Q) [21]. When evaluated in terms of domain information, other mutated residues in rs41382345 (E139V), rs121918394 (K164Q), rs11542041 (R132S) rs200703101 (R154P), rs121918393 (R154S), rs387906567 (R160C), rs7412 (R176C), rs11542040 (P102T) positions were reported that they are in an area which is important for the binding of molecules. Consequently, the mutation may disrupt the interaction with other molecules for binding and therefore affect the role of the protein. It has been reported that the mutated residue in the rs769455 (R163C) position is embedded in a domain which is significant for the binding. These differences may disrupt the structure of this domain and thus affect the binding features. The mutated residue at position rs11542029 (R50C) was reported to be located on the surface of an area of unknown function [21]. rs7412 and rs121918393 are located in receptor binding domain [22]. In terms of structure, mutations at positions E139V, P102T, K164Q, R132S, R154P, R154S, R160C, R163C, R176C are located in a sequence of residues annotated as a special region in UniProt: 8 X 22 AA approximate tandem repeats. It has been reported that differences in amino acid features may affect this area and impair its function [21]. In terms of variant information, mutations at positions K164Q, R154S, R160C, R163C, and R176C match the variants described previously, according to the ExPASy site at position K164Q (ExPASy site: VAR_00066), at position R154S (ExPASy site VAR_000656), at position R160C (VAR3000658) (ExPASy site: VAR_000659 and R176C variants (ExPASy site: VAR_000664) are reported to be significantly associated with disease grade [21].

Table 4 shows the information of hydrogen bonds and salt bridges formed by wild-type residues belonging to the R176C, R163C, R50C, R154S, R154P, R132S and E139V positions. Due to the size differences of amino acid residues, hydrogen bonds may not be formed as the original wild-type residue did. Similarly, ionic interactions will be disrupted due to amino acid change [21].

Table 1. Bioinformatical results of SNVs within the *APOE* gene

SNV ID	Amino acid change	SIFT	Score	PolyPhen-2 HumDiv	Score	PolyPhen-2 HumVar	Score	PROVEAN	Score	SNPs&GO	RI	Probability	Mutation Assessor	Score	SNAP 2	Score	Expected Accuracy
rs7412	R176C	D	0.001	PD	1000	PD	1.000	D	-3.936	Ds	0	0.520	M	2.28	E	56	75%
rs769455	R163C	D	0.008	PD	1.000	PD	0.999	D	-4.912	Ds	0	0.502	M	2.215	E	43	71%
rs11542029	R50C	D	0	PD	1.000	PD	1.000	D	-6.003	Ds	8	0.881	M	3.185	E	49	71%
rs121918393	R154S	D	0.007	PD	1.000	PD	0.997	D	-4.508	Ds	5	0.759	M	2.19	E	67	80%
rs121918394	K164Q	D	0.036	PD	0.996	PsD	0.994	D	-3.181	Ds	2	0.578	M	2.28	E	23	63%
rs200703101	R154P	D	0.007	PD	1.000	PD	0.999	D	-5.207	Ds	6	0.814	M	2.19	E	69	80%
rs387906567	R160C	D	0.001	PD	1.000	PD	1.000	D	-5.755	Ds	0	0.520	M	2.31	E	37	66%
rs11542040	P102T	D	0	PD	0.978	PsD	0.803	D	-6.648	Ds	5	0.749	M	2.28	E	52	75%
rs11542041	R132S	D	0.005	PD	1.000	PD	1.000	D	-4.213	Ds	7	0.841	M	2.31	E	63	80%
rs41382345	E139V	D	0.002	PD	0.999	PD	0.986	D	-5.993	Ds	7	0.867	M	2.28	E	51	75%

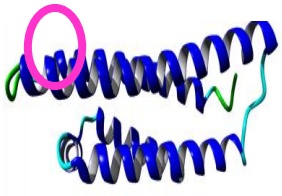
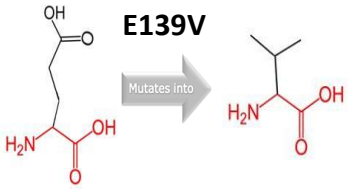
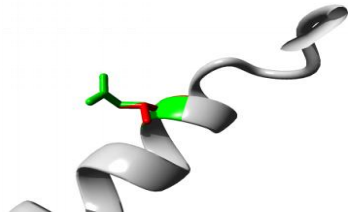
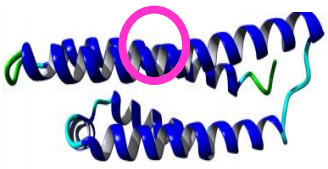
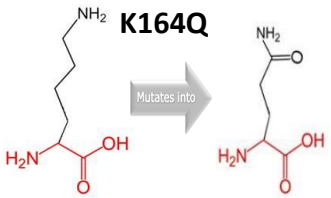
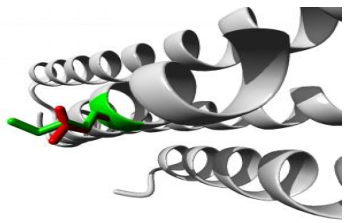
*D: Deleterious, PD: Probably Damaging, PsD: Possibly Damaging, Ds: Disease, RI: Reliability Index, Pr: Probability, M: Medium, E: Effect

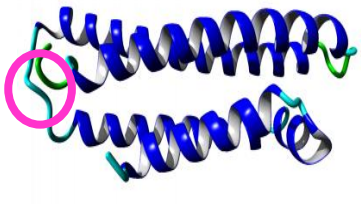
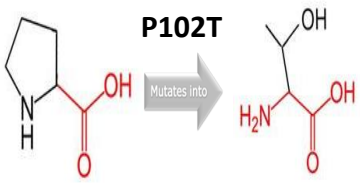
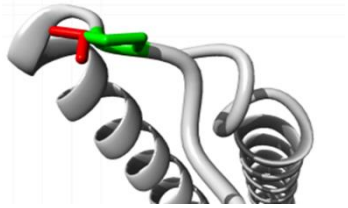
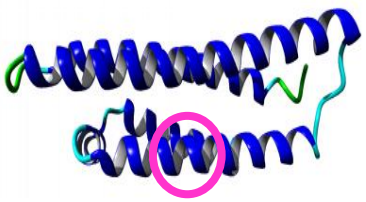
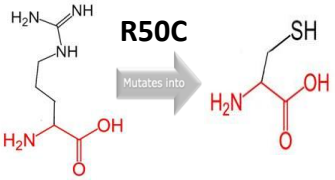
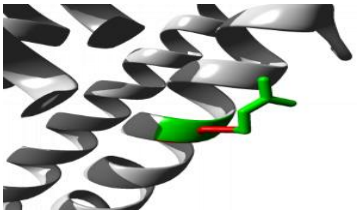
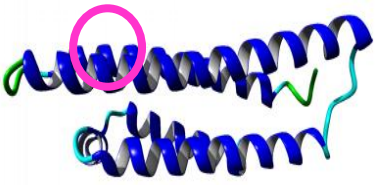
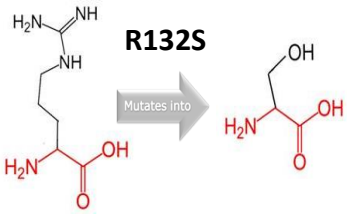
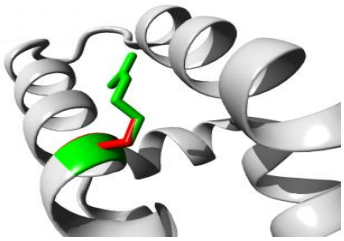
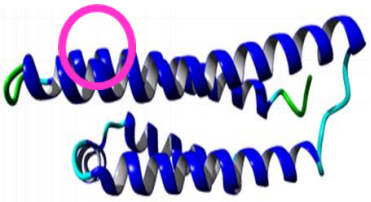
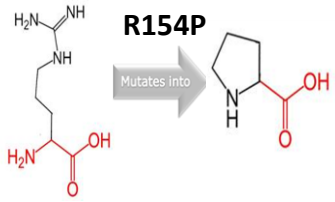
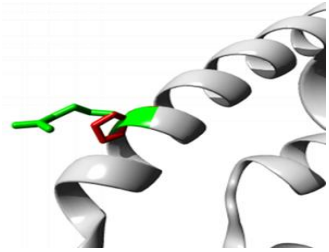
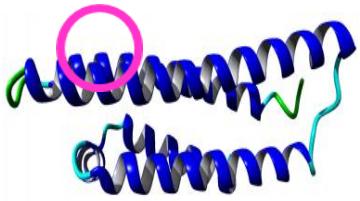
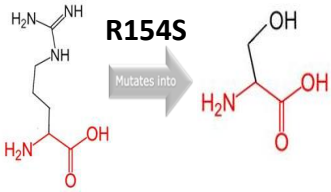
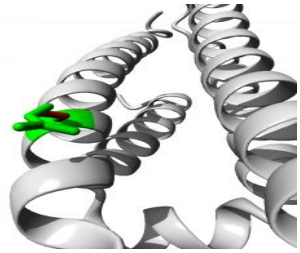
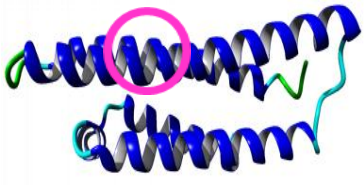
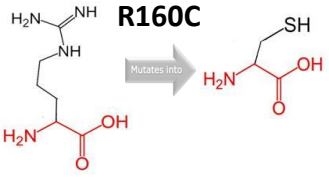
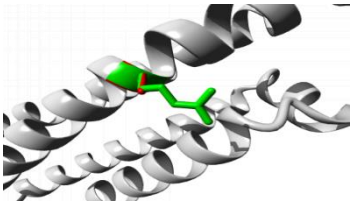
Table 2. Stabilization results of amino acid changes within APOE

SNV ID	Amino acid change	I-Mutant 3.0 Result	RI	MUpro Result	MUpro DDG
rs7412	R176C	Dec	6	Dec	0.877
rs769455	R163C	Dec	5	Dec	-0.760
rs11542029	R50C	Dec	6	Dec	-1.311
rs121918393	R154S	Dec	8	Dec	-1.043
rs121918394	K164Q	Inc	2	Dec	-0.771
rs200703101	R154P	Dec	3	Dec	-1.157
rs387906567	R160C	Dec	4	Dec	-0.565
rs11542040	P102T	Dec	7	Dec	-0.930
rs11542041	R132S	Dec	8	Dec	-0.841
rs41382345	E139V	Inc	3	Dec	-0.288

***** DDG: Delta Delta G, RI: Reliability Index, Dec: Decrease, Inc: Increase

Table 3. Project HOPE modeling results of harmful SNVs in the APOE gene

The protein in ribbon-presentation	Amino acid exchange	Close-up of the mutation
		
		

	 P102T Mutates into	
	 R50C Mutates into	
	 R132S Mutates into	
	 R154P Mutates into	
	 R154S Mutates into	
	 R160C Mutates into	

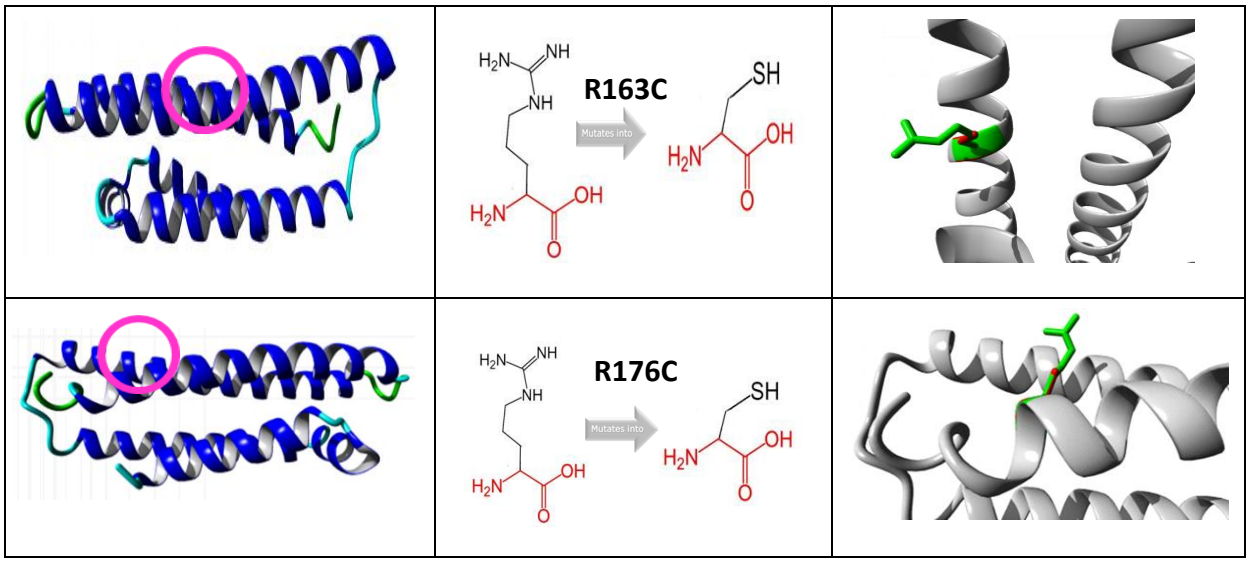


Table 4. Connections of wild type residues from Project HOPE

Amino acid change	**Hydrogen bond**	**Salt Bridge**
R176C	Asp position 172	Glu position 114 Asp position 172
R163C	Gln position 59	
R50C	Glu position 84	Glu position 84 Glu position 88
R154S	Glu position 150	Glu position 139 Glu position 150
K164Q	-	-
R154P	Glu position 150	Glu position 139 Glu position 150
R160C	-	-
P102T	-	-
R132S	His position 158	Glu position 139
E139V	Gln position 135	Arg position 132 Arg position 154

4. Discussion

APOE genotypes were reported to be associated with dysbetalipoproteinemia and hyperlipidemia [7,23]. In a study conducted in 2010, the association between 22 polymorphisms and four serum lipids in three main racial groups in the United States were investigated. rs7412 (R176C) and haplotype T-T were found to be strongly associated with LDL cholesterol and total cholesterol in non-hispanic blacks [24]. Furthermore, rs7412

polymorphism is reported as “landmark contra and pro risk factors” for AD [25]. Al-Eitan et al. investigated whether there is any relationship between rs2108622 of CYP4F2, rs7412, and rs405509 of APOE, and rs1801272 of CYP2A6 and dose variability of cardiovascular disease (CVD) and warfarin. Only APOE rs7412 was found to be associated with the CVD risk of the four selected SNPs [26]. In another study investigating the association of APOE hypermethylation with coronary heart disease (CHD) in men, APOE methylation levels were determined by PCR in 640 CHD cases and 436 controls, and genotyping of APOE rs7412 and rs7259620 was performed. As a result of the study, it was reported that APOE methylation is important for rs7412 and rs7259620 in terms of CHD risk in men [27]. Similarly, in our study, it was found that rs7412 (R176C) has a possibly harmful effect. In another study, it was reported that APOE gene polymorphisms were associated with CVD and defined the APOE ε4 allele as an independent risk factor for both Type 2 diabetes and CVD [28]. Saeed et al. (2018) were investigated SNVs with detrimental effects on the APOE different databases. They reported that rs769455 (R163C) was predicted to be deleterious [23]. Abdalla et al. (2017) were evaluated the SNVs in the APOE e4 gene. Among the 10 deleterious mutations found rs11542029 (R50C), rs200703101 (R154P), rs11542041 (R132S) and rs41382345 (E139V) are similar to our study [11]. In a previous study, SNVs in the *APOE* gene were examined via computational tools. Among eighty-eight SNVs analyzed, rs11542041 (R132S) was determined to have a significant effect on functional variation [10]. In a study conducted in 2021 on the APOE gene, which is also known as a candidate gene for autosomal dominant hypercholesterolemia, they reported that three variants (p.Arg154Ser, p.Leu167del, p.Glu230Lys) were found in families with dominant transmission of the mixed hyperlipidemia phenotype defining familial combined hyperlipidemia (FCHL) [9]. In addition, Halil et al. reported some variations in the APOE gene associated with dyslipoproteinemia in their study. rs7412 (R176C), rs769455 (R163C), rs121918393 (R154S), rs121918394 (K164Q) and rs11542041 (R132S) variants were also found to have potentially harmful effects in our study [9]. Masoodi et al. reported that rs11542041 (R132S), rs11542040 (P102T) and rs11542034 (E150G) were to be potentially functional SNVs via bioinformatic analysis [8].

Amino acids show differences in charge, size, and hydrophobicity. The size differences of mutant and wild-type amino acids can lead to disruption of the interaction of the protein with other molecules or to the deterioration of its structure by creating a wrong conformation with another residue. [21]. Hydrophobic or hydrophilic effects have an effect on protein folding and nucleic acid stability. For a protein to function efficiently, it must be folded correctly. Variations with polymorphism can cause loss of hydrophobic interactions in the core or surface of the protein [29,30]. The charge difference that may occur between mutant and wild type residues can disrupt ionic interactions [31]. In our study, the wild-type residue is larger than the mutant residue in the variations of R176C, R163C, R50C, R154S, K164Q, R154P, R160C, R132S, and E139V. These variations can cause a loss of external interaction. The mutant residue was more hydrophobic than the wild-type residue in the variations of R176C, R163C, R50C, R154S, R154P, R160C, R132S, E139V. This difference in hydrophobicity may impair the correct folding of the protein. The wild-type residue was positive in R176C, R163C, R50C, R154S, K164Q, R154P, R160C, R132S variants, while mutant residue was neutral. Charge differences can disrupt ionic interactions. In addition, the wild-type amino acid can be found at an important site for the binding of other molecules. With the variation that occurs, the domain structure may deteriorate, which may affect the binding properties and thus the protein structure. Variations in E139V, P102T, K164Q, R132S, R154P, R154S, R160C, R163C, and R176C positions are in UniProt: 8 X 22 AA tandem repeat area. The resulting variations may cause dysfunction in the repetitive area. [21].

As a result, either experimental or computational studies have been found in the literature about rs7412, rs769455, rs11542029, rs121918393, rs121918394, rs200703101, rs387906567, rs11542040, rs11542041 and rs41382345 which are predicted to have a high-risk effect within the *APOE* gene in this study. Our analysis should be valuable for further proteomic, genomic, and clinical research because of the associations between SNVs within the APOE gene and various diseases such as Alzheimer's disease etc.

References

[1] Zuo L, van Dyck CH, Luo X, Kranzler HR, Yang BZ, Gelernter J. Variation at APOE and STH loci and Alzheimer's disease. **Behavioral and Brain Functions,** 2(1),1-10, (2006).

[2] Liu CC, Kanekiyo T, Xu H BG. Apolipoprotein E and Alzheimer disease: risk, mechanisms, and therapy. **Nature Reviews Neurology,** 9:106–18, (2013).

[3] Mahley RW, Huang Y, Rall SC. Pathogenesis of type III hyperlipoproteinemia (dysbetalipoproteinemia): Questions, quandaries, and paradoxes. **Journal of lipid research,** [Internet];40:1933–49, (1999). Available from: http://dx.doi.org/10.1016/S0022-2275(20)32417-2

[4] Gkouskou K, Vasilogiannakopoulou T, Andreakos E, Davanos N, Gazouli M, Sanoudou D, et al. COVID-19 enters the expanding network of apolipoprotein E4-related pathologies. **Redox biology,** 41, 101938, (2021).

[5] Finch CE, Kulminski AM. The ApoE Locus and COVID-19: Are we going where we have been? Journals Gerontol. - **The Journals of Gerontology: Series A,** 76(2), e1-e3, (2021).

[6] Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GW, et al. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. **Science,** 261(5123), 921-923, (1993).

[7] Solanas-Barca M, de Castro-Orós I, Mateo-Gallego R, Cofán M, Plana N, Puzo J, et al. Apolipoprotein E gene mutations in subjects with mixed hyperlipidemia and a clinical diagnosis of familial combined hyperlipidemia. **Atherosclerosis,** [Internet];222:449–55, (2012). Available from: http://dx.doi.org/10.1016/j.atherosclerosis.2012.03.011

[8] Masoodi TA, Al Shammari SA, Al-Muammar MN, Alhamdan AA. Screening and evaluation of deleterious SNVs in APOE gene of Alzheimer's disease. **Neurology research international,** 2012, (2012).

[9] Khalil YA, Rabès JP, Boileau C, Varret M. APOE gene variants in primary dyslipidemia. **Atherosclerosis,** 328:11–22, (2021).

[10] Namboori PKK, Vineeth K V., Rohith V, Hassan I, Sekhar L, Sekhar A, et al. The ApoE gene of Alzheimer's disease (AD). **Functional & İntegrative Genomics,** 11(4), 519–22, (2011).

[11] AA Abdalla A. In Silico Analysis of Single Nucleotide Polymorphisms (SNVS) in Human Abetalipoprotein Epsilon 4 (APOE E4) Gene as a Cause of Alzheimer's Disease from Genetic Mutation to Functional Predication from Structural Change Patterns. **Current Trends in Biomedical Engineering & Biosciences,** 5:40–2, (2017).

[12] Kaman T, Karasakal ÖF, Özkan Oktay E, Ulucan K, Konuk M. In silico approach to the analysis of SNVs in the human APAF1 gene. **Turkish Journal of Biology,** 43(6), 371-381, (2019).

[13] Özkan Oktay E, Kaman T, Karasakal ÖF, Ulucan K, Konuk M, Tarhan N. Alzheimer Hastalığı ile İlişkilendirilen APH1A Genindeki Zararlı SNV'lerin In Silico Yöntemler

ile Belirlenmesi. S**üleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi**, 23(2), (2019).

[14] Vaser R, Adusumalli S, Leng SN, Sikic M, Ng PC. SIFT missense predictions for genomes. **Nature protocols,** [Internet]; 11:1–9, (2016). Available from: http://dx.doi.org/10.1038/nprot.2015-123

[15] Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. **Current protocols in human genetics,** 76(1), 7-20, (2013).

[16] Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the Functional Effect of Amino Acid Substitutions and Indels. **PLoS One**;7, (2012).

[17] Reva B, Antipin Y, Sander C. Determinants of protein function revealed by combinatorial entropy optimization. **Genome biology,** 8(11), 1-15, (2007).

[18] Calabrese R, Capriotti E, Fariselli P, Martelli PL, Casadio R. Functional annotations improve the predictive score of human disease-related mutations in proteins. **Human mutation,** 30:1237–44, (2009).

[19] Capriotti E, Calabrese R, Casadio R. Predicting the insurgence of human genetic diseases associated to single point protein mutations with support vector machines and evolutionary information. **Bioinformatics,** 22:2729–34, (2006).

[20] Cheng J, Randall A, Baldi P. Prediction of protein stability changes for single-site mutations using support vector machines. **Proteins: Structure, Function, and Bioinformatics,** 62:1125–32, (2006).

[21] Venselaar H, te Beek TAH, Kuipers RKP, Hekkelman ML, Vriend G. Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. **BMC Bioinformatics,** 11(1), 1-10, (2010).

[22] Rasmussen KL, Tybjærg-Hansen A, Nordestgaard BG, Frikke-Schmidt R. APOE and dementia – resequencing and genotyping in 105,597 individuals. **Alzheimer's & Dementia,** 16:1624–37, (2020).

[23] Saeed NAHAAH, Hamzah IH, Ali ANM, Abuderman AA. Prediction of single nucleotide polymorphisms (SNVs) in apolipoprotein E gene and their possible associations with a deleterious effect on the structure and functional properties: an in silico approach. **Network Modeling Analysis in Health Informatics and Bioinformatics/s,** [Internet]; 7:1–6, (2018). Available from: http://dx.doi.org/10.1007/s13721-018-0178-9

[24] Chang M huei, Yesupriya A, Ned RM, Mueller PW, Dowling NF. Genetic variants associated with fasting blood lipids in the U.S. population: Third National Health and Nutrition Examination Survey. **BMC Medical Genetic,** 11(1), 1-13, (2010).

[25] Kulminski AM, Shu L, Loika Y, He L, Nazarian A, Arbeev K, et al. Genetic and regulatory architecture of Alzheimer's disease in the APOE region. **Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring,** 12(1), e12008, (2020).

[26] Al-Eitan, L. N., Almasri, A. Y., Alnaamneh, A. H., Aman, H. A., Alrabadi, N. N., Khasawneh, R. H., & Alghamdi, M. A. Influence of CYP4F2, ApoE, and CYP2A6 gene polymorphisms on the variability of Warfarin dosage requirements and susceptibility to cardiovascular disease in Jordan. **International journal of medical sciences,** 18(3), 826, (2021).

[27] Ji, H., Zhou, C., Pan, R., Han, L., Chen, W., Xu, X., ... & Duan, S. APOE hypermethylation is significantly associated with coronary heart disease in males. ***Gene,*** *689*, 84-89, (2019).

[28] El-Lebedy, D., Raslan, H. M., & Mohammed, A. M. Apolipoprotein E gene polymorphism and risk of type 2 diabetes and cardiovascular disease. **Cardiovascular diabetology,** 15(1), 1-11, (2016).

[29] Damodharan, L., & Pattabhi, V. Hydropathy analysis to correlate structure and function of proteins. **Biochemical and biophysical research communications,** 323(3), 996-1002, (2004).

[30] Atalay, S. In Silico Analysis of the Structural and Functional Consequences of Polymorphic Amino Acid Substitutions in the Cattle HSF1 Protein. **Kafkas Üniversitesi Veteriner Fakültesi Dergisi,** 28(3), (2022).

[31] Biro, J. C. Amino acid size, charge, hydropathy indices and matrices for protein structure analysis. **Theoretical Biology and Medical Modelling,** 3(1), 1-12, (2006).