REPUBLIC OF TURKEY
ACIBADEM MEHMET ALİ AYDINLAR UNIVERSITY
INSTITUTE OF HEALTH SCIENCES

TARGETED NEXT-GENERATION SEQUENCING OF BASE EXCISION REPAIR GENES FOR LATE-ONSET ALZHEIMER'S AND PARKINSON DISEASE'S RISK

TUĞÇE ERTÜZÜN

DOCTORATE THESIS

DEPARTMENT of MEDICAL BIOTECHNOLOGY

SUPERVISOR
Prof. Dr. Meltem MÜFTÜOĞLU

ISTANBUL-2020

REPUBLIC OF TURKEY
ACIBADEM MEHMET ALİ AYDINLAR UNIVERSITY
INSTITUTE OF HEALTH SCIENCES

TARGETED NEXT-GENERATION SEQUENCING OF BASE EXCISION REPAIR GENES FOR LATE-ONSET ALZHEIMER'S AND PARKINSON DISEASE'S RISK

TUĞÇE ERTÜZÜN

DOCTORATE THESIS

DEPARTMENT of MEDICAL BIOTECHNOLOGY

SUPERVISOR
Prof. Dr. Meltem MÜFTÜOĞLU

ISTANBUL-2020

DECLARATION

I declare that this thesis study is my own study, did not have an unethical behavior at all stages from planning of the thesis to writing, procured all the information in this thesis within academic and ethical rules, showed references to all information and comments that has not conducted with this thesis study,and included these references to the references list and also have not violated patent and copyrights during this thesis study and writing.

21.07.2020

Tuğçe ERTÜZÜN

ACKNOWLEDGEMENTS

First of all, I would like to thank Prof. Dr. Meltem Müftüoğlu, for making this project come to life and involving me, and for guiding me through all her knowledge and experience at every stage of the study. Also, I would like to thank Prof. Dr. Uğur Sezerman, Assoc. Prof. Nadja de Souza Pinto and department of Neurology of Medeniyet University Goztepe Training and Research Hospital for their contributions to the project.

This PhD study was supported by TUBITAK Scientific and Technologic Research Project Supporting Program with 114Z875 grant number. I would like to thank TUBITAK for their support and to Acıbadem Mehmet Ali Aydınlar Univeristy for all the facilities to complete this study.

I would like to express my thanks to my lab mates. Especially to the ones who become my friends, Berna Somuncu who I can always rely on and Ayşegül Ekmekçioğlu, who was an excellent copilot in this study, for always staying rock-solid for me.

Many thanks to my friends who stayed with me both in my bad and good days and don't make me feel alone.

To my mother, who has given me all her support emotionally and for studying a complete new field just to make her daughter more comfortable, to Yılmaz, who has been there for me without getting tired from the beginning to the end and to Fındık and Gece for bringing light to my darkness, I would like to thank especially. Last but not least, many thanks to my all family members who supported me through all the ups and downs of this journey and always respected and supports all my decisions.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND SYMBOLS

Abbreviation	Explanation
5-OHU	5-hydroxyuracil
8-OHGua	8-hydroxyguanine
AD	Alzheimer's Disease
AP	Amyloid plaque
APE1	AP endonuclease 1
APOE	Apoliprotein E
APP	Amyloid precursor protein
ATP13A2	ATPase Cation Transporting 13A2
Aβ	Amyloid β
βA40	40 aminoacid length Amyloid β
βA42	42 aminoacid length Amyloid β
BER	Base excision repair
CD	Cerebrovascular disease
CDR	Clinical Dementia Rating
DNAJC13	DnaJ heat shock protein family (Hsp40) member C13
DNAJC6	DnaJ Heat Shock Protein Family (Hsp40) Member C6
dRP	5'-deoxyribose phosphate
EIF4G1	Eukaryotic translation initiation factor 4 gamma 1
EOAD	Early-onset Alzheimer's disease
ε2	APOE ε2 allele
ε3	APOE ε3 allele
ε4	APOE ε4 allele
Fapy	Formamidopyrimidine
FapyAde	4,6-diamino-5-formamidopyrimidine
FapyGua	2,6-diamino-4-hydroxy-5-formamidopyrimidine
FBXO7	F-box protein 7
FTLD	Frontotemporal lobar degeneration
GBA	Glucosylceramidase beta
GWAS	Genome wide association study
HTRA2	HtrA serine peptidase 2
hpC	High-pathology control
LBD	Lewy body dementia
LD	Linkage disequilibrium
LDL	Low density lipoprotein
LIG1	DNA ligase I
LIGIIIα	DNA ligase IIIα
LOAD	Late-onset Alzheimer's disease
LOD	Logarithm of odds
LP-BER	Long patch base excision repair

LRRK2	Leucine rich repeat kinase 2
MA	Minor allele
MAF	Minor allele frequency
MCI	Mild cognitive impairment
mtDNA	Mitochondrial DNA
NEIL1	Endonuclease VIII-like DNA glycosylase 1
NFT	Neurofibrillary tangle
NGS	Next generation sequencing
OGG1	8-oxoguanine DNA glycosylase
OR	Odds ratio
PARK7	Parkinsonism associated deglycase
PD	Parkinson's disease
PINK1	PTEN-induced kinase 1
PLA2G6	Phospholipase A2 Group VI
PNKP	Polynucleotide kinase-phosphatase
Polγ	DNA polymerase γ
POLβ	DNA polymerase β
POLβ	DNA polymerase beta
POLγ	DNA polymerase gamma
PRKN	Parkin RBR E3 Ubiquitin Protein Ligase
SNCA	Aynuclein alpha
SP-BER	Short patch base excision repair
SSB	Single strand break
SYNJ1	Synaptojanin 1
TBI	Traumatic brain injury
TMEM230	Transmembrane protein 230
UCHL1	Ubiquitin carboxy-terminal hydrolase L1
UNG	Uracil DNA glycosylase
UNG	Human Uracil DNA Glycosylase
VPS13C	Vacuolar Protein Sorting 13 Homolog C
VPS35	VPS35 Retromer Complex Component
WHO	World Health Organization
wt	Wild type
XRCC1	X-Ray repair cross complementing 1

LIST OF FIGURES

LIST OF TABLES

SUMMARY

Late-onset Alzheimer's disease (LOAD) is a multifactorial neurodegenerative disorder having the highest prevalence. The roles of base excision repair (BER) deficiency and accumulation of oxidative DNA damage on LOAD etiology and progression are known. However, it is not known whether genetic variant(s) of specific BER genes contribute to reduced BER activity in LOAD patients and whether they are associated with risk, development and/or progression of LOAD. Therefore in this thesis study, promoter, exon and intron regions of human uracil glycosylase (*UNG*), human endonuclease VIII-like DNA glycosylase 1 (*NEIL1*) and human DNA polymerase β (*POLβ*) were sequenced using targeted next-generation sequencing. Blood samples of Turkish LOAD patients, age-matched controls and temporal cortex (TC) and cerebellum (CE) post-mortem tissues of Brazilian LOAD patiens, high-pathology controls (hpCs) were included to this study. Also, sporadic Parkinson's disease (PD) patients' blood samples were included to this study. Apoliprotein E (*APOE*), the known LOAD genetic risk factor was also sequenced to study the association of BER gene variants and *APOE*. In Turkish LOAD patients' blood samples *UNG* rs1610925, rs2268406, rs80001089, rs1018782 and rs1018783 variants and in LOAD-TC only *UNG* rs80001089 are found to be statistically significantly related. Significant *UNG* rs80001089 and *NEIL1* rs7182283 epistasis is found only in LOAD TC samples. *APOE* rs769446 is the only significant PD related variant. These results suggest, significant BER gene variants may be associated with the risk of LOAD in non-APOE ε4 carrier.

Key words: Base Excision Repair, Gene Variant, Late-onset Alzheimer's Disease, Next Generation Sequencing, Sporadic Parkinson's Disease

ÖZET

Geç Başlangıçlı Alzheimer ve Parkinson Hastalığı Riski için Baz Eksizyon Tamir Genlerinin Hedeflenmiş Yeni Nesil Dizilemesi

Geç başlangıçlı Alzheimer hastalığı (LOAD) en yüksek prevelanslı multifaktöriyel nörodejeneratif bir hastalıktır. Oksidatif DNA hasarı ve baz eksizyon tamiri (BER) bozukluklarının LOAD etiyolojisi ve ilerlemesi üzerindeki rolleri bilinmektedir. Bununla birlikte, belirli BER genlerinin genetik varyantlarının LOAD hastalarında BER aktivitesinin azalmasına katkıda bulunup bulunmadığı ve LOAD'in risk, gelişme ve / veya ilerlemesi ile ilişkili olup olmadığı bilinmemektedir. Bu nedenle bu tez çalışmasında, insan urasil glikosilaz (*UNG*), insan endonükleaz VIII-benzeri DNA glikosilaz 1 (*NEIL1*) ve insan DNA polimeraz β (*POLβ*) 'nın promotörü, ekzon ve intron bölgeleri, hedeflenmiş yeni nesil sekanslama kullanılarak dizilendi. Türk LOAD hastaları ve yaş uyumlu kontrollerinin kan örnekleri, Brezilya LOAD hastalarının, yaş uyumlu ve yüksek patolojili kontrollerinin, temporal korteks (TC) ve beyincik (CE) post mortem dokuları çalışmaya dahil edildi. Ayrıca, sporadik Parkinson hastalığı (PD) hastalarının kan örnekleri de çalışmaya dahil edildi. Bilinen LOAD genetik risk faktörü olan Apoliprotein E (*APOE*) de BER gen varyantları ve APOE'nin birlikteliğini incelemek üzere sekanslandı. Türk LOAD hastanın kan örneklerinde rs1610925, rs2268406, rs80001089, rs1018782 ve rs1018783 varyantlarının ve LOAD-TC'de sadece *UNG* rs80001089'un istatistiksel olarak anlamlı derecede ilişkili olduğu bulunmuştur. Anlamlı *UNG* rs80001089 ve *NEIL1* rs7182283 epistazı sadece LOAD TC örneklerinde rastlannıştur. *APOE* rs769446, PD ile ilişkili tek anlamlı varyanttır. Bu sonuçlar, APOE ε4 taşıyıcısı olmayanlarda önemli BER gen varyantlarının, LOAD riski ile ilişkili olabileceğini düşündürmektedir.

Anahtar kelimeler: Baz Eksizyon Tamiri, Geç Başlangıçlı Alzheimer Hastalığı, Gen Varyantı, Sporadik Parkinson Hastalığı, Yeni Nesil Dizileme

1. BACKGROUND AND AIM OF THE STUDY

Late-onset Alzheimer's disease (LOAD) is a multifactorial neurodegenerative disorder having the highest prevalence. LOAD, which is the most common type of dementia that affects large areas of the cerebral cortex and hippocampus, primarily manifests as losing memory and reasoning ability and developing comprehension problems in the patient. LOAD has two distinguishing histopathological markers, amyloid plaques (AP) and neurofibrillary tangles (NFTs). LOAD is seen in 95-99% of AD patients. The main risk factors are; family history, age and genetic risk factors but the etiology has not fully understood yet (1,2).

Sporadic Parkinson's disease (PD) is characterized by loss of neurons in black matter and accumulation of α-synuclein protein. The disappearance of dopaminergic neurons causes the reduction of dopamine involved in brain-muscle coordination. Sporadic PD patients experience tremors of hands, arms, legs, jaw and face, stiffness in limbs and trunk, slowing movements and difficulties in balance, speech and coordination. The severity of these symptoms increases with the degree of the disease (3,4).

Studies in post-mortem brain tissues of LOAD patients in the preclinical and late stages, peripheral blood samples, mouse models, and cell lines revealed high oxidative DNA damage and low BER activity (5-8). Like LOAD, 90% of patients diagnosed with PD are sporadic PD patients, and their etiology has not been fully elucidated. Studies in post-mortem brain tissues, mouse models, and cell culture of sporadic PD patients have revealed the association of sporadic PD with oxidative stress (9-12).

Stimulation of neuronal apoptosis in *UNG* gene-silenced mice reveals the importance of UNG protein in neuronal development. The activity and protein amount of UNG protein were found to be low in LOAD brain tissues. UNG protein is uracil-

specific and low UNG activity results in uracil accumulation. Accumulating uracil makes neurons more prone to amyloid inhibitory protein toxicity and stimulates neuronal apoptosis (13-15). Studies with *NEIL1* gene-silenced mice have demonstrated that *NEIL1* plays an important role in short- and long-term memory loss and cognitive decline. Studies in LOAD brain tissues have demonstrated that the level and activity of NEIL1 protein is significantly low, and a decrease in the *NEIL1* gene expression level in LOAD (16-19). In studies with silenced POLβ gene in mice, embryonic deaths have occurred in the late stages of embryogenesis, which has been associated with the neuronal developmental disorder. In experiments with mice and human neuron cells, it has been shown that *POLβ* disorder causes neuronal cell death, impaired central nervous system development, and become more susceptible to oxidative stress. In studies conducted with the 3xTg 3xTg AD/Polβ$^{+/-}$ mouse model, it has been reported that neurodegeneration is increased and the AD phenotype is observed (20,21). No direct relationship was found between *UNG* or *NEIL1* and PD. Studies in cell culture and animal models with sporadic PD have been associated with increased loss of POLβ dopaminergic neurons (22).

Base excision repair (BER) defects and concomitant oxidative DNA damage accumulation play a role in the etiology and progression of late-onset Alzheimer's disease (LOAD). In this thesis study, to find out if genetic variants of three BER genes, *UNG*, *NEIL1* and *POLβ*, are responsible for the decreased BER mechanism in LOAD patients, and if the variants of these genes are related to risk, development and progression of LOAD, and also sporadic PD, *UNG*, *NEIL1* and *POLβ* genes, were sequenced for the first time with the new generation sequencing method, including promoter, intron and exon regions.

2. INTRODUCTION

2.1. Neurodegenerative Disorders

One of the most frequent disease in elderly is dementia and worldwide it is 3rd cause of death. “World Health Organization (WHO)”, declares that dementia is a syndrome that affects memory, other cognitive abilities and behavior by interfering daily living activity abilities. In 2017, around 50 million people was affected from dementia and it is estimated that the number 152 million in 2050 (23). In Turkey, a study by Ministry of Family and Social Policies performed in 2017 estimated that 600.000 people are facing with dementia. The study also pointed out that half of the patients cannot be diagnosed therefore the estimation may not be representing the actual number (24). Dementias can be grouped according to their treatability. Reversible dementias can be caused by drugs, alcohol or metabolic disorders while dementias caused by the neurodegenerative disorders are irreversible. Neurodegeneration is characterized with progressive loss of neurons. Apart from the normal aging, for dementia age is the greatest risk factor. Memory loss is the most common symptom in neurodegenerative disorders, but it alone does not point out the disease. Two or more brain function deficiencies such as memory loss, decrease in processing speed, language skills, executive function problems etc. are seen together in neurodegenerative disorders (23,25).

Dementia patient’s 60 to 80% has Alzheimer’s disease (AD). It is the most frequently seen neurodegenerative disorder. But recent autopsy studies using AD patients’ post-mortem brain tissues showed that among 50% of patient AD pathology is not seen alone, instead, cerebrovascular disease (CD) or Lewy body dementia (LBD) pathologies can be seen. Another type of neurodegenerative disorder is Parkinson’s disease (PD), which seen in 3-4% of neurodegenerative disorder patients and mostly primarily characterized by problems with movement. Five to ten percent of the neurodegenerative disorder patients are CD patients. CD is seen in patients with damaged brain vessels and tissues which effects 5 to 10% of the dementia patients

alone. Usually CD is seen in mixed pathology with other dementia types. Protein alpha-synuclein's abnormal aggregations result as LBD. Like CD it is seen in 5 to 10% of the neurodegenerative disorder patients alone but also like CD it can be seen in mixed pathology with other dementia types. Atrophy of frontal and temporal lobes are the hallmark of frontotemporal lobar degeneration (FTLD). Less than 10% of the neurodegenerative disorder dementia patients have FTLD. In CD and FTLD symptoms are seen in earlier ages than 65 years but their symptoms are similar to AD (26).

Mini-mental state exam (MMSE) is a tool for determining the cognitive situation of the dementia patients. A series of questions are addressed to the patients to measure their everyday mental skills. The individuals who have scores between 20 and 24 are considered as mild dementia. Thirteen to 20 indicates moderate and lesser than 12 indicates severe dementia (27). Clinical Dementia Rating (CDR) is also another staging tool used for cognitive situation. It is calculated by six different main topics which are as follows; orientation, judgment, problem solving, memory, personal care, community affairs and home and hobbies performance. There are 5 different stages in CDR scoring; none (0), questionable (0.5), mild (1), moderate (2) and severe (3) (28).

2.1.1. Alzheimer's disease

AD is a progressive and multifactorial neurodegenerative disorder. AD is characterized by memory, mobility and coordination deprivation, cognitive decline, and various neuropsychiatric symptoms. AD stages are divided into three major groups: mild, moderate and severe. Mild stage of AD is the most difficult stage for diagnosis because patients suffer from symptoms that can be confounded with forgetfulness that is normal part og aging. In the moderate stages, the problems of losing memory and reasoning and comprehension are seen, while in the severe stages the magnitude of memory loss increase and patients develop symptoms such as paranoia and delusions. Towards the severe stages, patients lose their ability to communicate completely and become in need of care. Frontal and temporal lobes are the first affected regions of AD. The progression of the disease through the other regions of the brain show differences among patients (29).

As Braak and Braak stated in 1991, the staging according to amyloid plaques (AP) and neurofibrillary tangle (NFT) distribution differs from each other. They staged the AD for AP as stages A, B and C and for NFT as stages I-II, III-IV and V-VI. The difference between the stages is mainly the localization of APs and NFTs. In AP staging A, the amyloid deposits localized mainly in isocortex and frontal, temporal, and occipital lobes basal portions. In stage B, in all isocortical association areas except primary motor field and sensory areas APs can be seen. APs are seen in all isocortical areas as densely packed deposits. NFT stages I-II are characterized by NFT deposits in transentorhinal regions. In stages III-IV, NFTs are spreaded through entorhinal and transentorhinal layer and mildly to hippocampal and isocortical regions. Hallmark of the stages V-VI is severely affected isocortex. The distribution of NFTs from stage I to VI gradually increases (2). The bridging between clinical symptoms of AD and neuropathological hallmarks was started by Grober and colleagues in 1999. The AP stage A and NFT stages I-II are grouped as a mild, AP stage B and NFT stages III-IV are grouped as a moderate and finally AP stage C and NFT stages V-VI are grouped as a severe (30).

AP and NFTs are the two most distinct histopathological signatures of AD. Besides these signatures, synapse-neuron loss, gliosis, inflammation, losses in other neurotransmitter systems and a marked atrophy in the brain are also seen in AD. APs are made of amyloid β (Aβ) peptide. This peptide is a metabolism product of proteolytic cleavage of amyloid precursor protein (APP), with varying lengths (38-43 amino acid long). The neuronal function of APP has not been fully discovered yet. APP is enzymatically processed by amyloidogenic and non-amyloidogenic pathways. APP is processed by α-secretase enzyme to produce a soluble peptide in non-amyloidogenic pathway. But, in amyloidogenic pathway β and γ secretase enzymes process APP to produce 40 (βA40) and 42 (βA42) amino acid long peptides (31). βA42 deposits as diffuse plaques and forms APs. γ secretase complex is composed of four proteins; presenilin (PS), nicastrin (NCT), anterior pharynx-defective 1 (APH-1), and presenilin enhancer 2 (PEN2). Mutations on genes encoding γ secretase complex proteins are linked with AD related genetic mechanisms. These mutations are related with enhanced expression of βA42 and related β sheet deposition (32,33). NFTs are

made of TAU protein. This protein is a member of microtubule-associated proteins family. TAU protein is responsible from stabilization of axonal microtubules and, integrity of cytoskeleton and also has a responsibility in axonal transport. Hyperphosphorylated TAU via hyperactive kinases and hypoactive phosphatases loses its binding capacity to microtubules. Thus, it leads to polymerization of insoluble double stranded filaments. Studies using AD patient's post-mortem brain tissues and cerebrospinal fluid showed total and phosphorylated TAU increment. The mechanism underlying increased TAU protein has not been fully understood yet. Images of AP and NFTs are shown in Figure 2.1A. Figure 2-1B shows the regional positioning of APs and NFTs. The deposition severity of these histopathological signatures increases through the stages of AD (29).

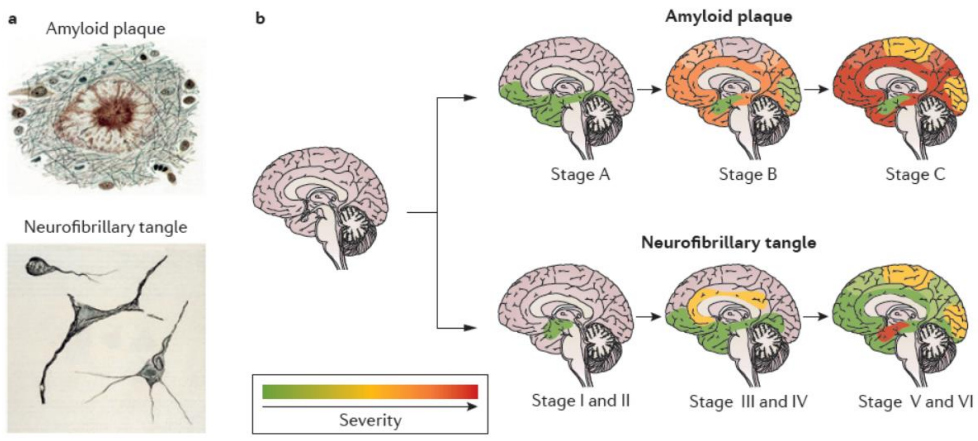

Figure 2.1. Alzheimer's Disease pathological development. A. Bielschowsky method of silver impregnation was used to visualize deposited proteins of APs and NFTs. B. Regional positioning of APs and NFTs through the stages of AD (29)

AD is catergorized as; early-onset AD (EOAD) and late-onset AD (LOAD). EOAD is mainly seen before the patient reaches to 65 years old and can be seen as early as age 30. Mostly AD patients are LOAD patients and the symptoms of these patients appear after 65 years of age (29).

EOAD is seen in the 1-5% percent of AD patients. *APP* gene and γ secretase complex (Presenilin-1 *(PSEN1)* and Presenilin-2 *(PSEN2))* are the definitive genetic risk factors for EOAD that related to increased levels of βA42. Sixty percent of the

EOAD patients have AD history in their family. Thirteen percent of family related EOAD disease is inherited in autosomal dominant manner. EOAD patients who carry *APP* and *PSEN1* gene mutations have a 100% penetrance and the disease can progress when the patient reaches 30 and may extend to age 60. *APP* gene is also a high EOAD genetic risk factor for individuals who have Down syndrome. Since it is coordinated on 21th chromosome, the trisomy of this chromosome is related to significant levels of APs and NFTs. These individuals 30% who are in their 50s have AD and in their 60s 50% has AD. *PSEN2* gene mutation carriers have 95% penetrance and the disease can develop when the patient reaches age 40-70 (34-36).

The etiology underlying the LOAD has not been fully discovered yet but the main risk factors are accepted as family history, genetic risk factors and age. Individuals who have family history of AD are found to have higher risk of having LOAD (37). Honea and colleagues showed that, patients with maternal LOAD history have significantly higher AD pathophysiology markers than patients with no maternal LOAD history (1). The genetic risk factors of LOAD are not definitive markers for LOAD but they are found to be as an enhancing factor of LOAD. The mutation of Apoliprotein E (*APOE*) gene, which is responsible from cholesterol and lipid transport, is the well-known and strongest risk factor of LOAD. On APOE gene 4th exon, 2 single nucleotide polymorphisms (SNPs) rs429358 and rs7412 result in 3 *APOE* alleles; ε2, ε3 and ε4. All humans carry one of the three alleles. Most frequent allele of APOE is ε3 (~80%) among humans, followed by ε4 (~14%) and ε2 (~8%). ε4 allele found to be dramatically increased in AD patients. But only 40-60% of the LOAD patients are ε4 carriers and the relation between *APOE* and LOAD is not clearly understood yet (38-40). Lately, genome-wide association studies (GWAS) linked LOAD and other genes such as; sortilin-related receptor-1 (*SORL1*), bridging integrator 1 (*BIN1*), complement component (3b/4b) receptor 1 (*CR1*), clusterin (*CLU*), phosphatidylinositol binding clathrin assembly protein (*PICALM*), ATP-binding cassette subfamily A member 7 (*ABCA7*), CD33, fermitin family member 2 (*FERMT2*), cas scaffold protein family member 4 (CASS4), membrane-spanning 4-domains, subfamily A (*MS4A6A*), EPH receptor 1 (*EPHA1*), histocompatibility complex class II, DRβ1 and 5, (*HLA-DRB1*, *HLA-DRB5*), protein tyrosine kinase 2

beta (*PTK2B*), CD2 associated protein (*CD2AP*), zinc finger CW-type and PWWP domain containing 1 (*ZCWPW1*), solute carrier family 24 (sodium/potassium/calcium exchanger), member 4/ras and rab interactor 3 (*SLC24A4/RIN3*), inositol-polyphosphate 5-phosphatase (*IN5PPD*), myocyte enhancer Factor 2C (*MEF2C*), NME/NM23 family member 8 (*NME8*), CUGBP Elav-like family member 1 (*CELF1*) locus, CD33 molecule (*CD33*). These new genetic risk factors for LOAD have lower odds ratio (OR) when compared with *APOE* gene ε4 allele (14-46). Besides these major risk factors, cardiovascular diseases, raumatic brain injury (TBI), schooling, lack of social and intellectual responsibilities are accepted as modifiable risk factors for LOAD (47).

A recent subgroup of AD was emerged when Iacono and colleagues named individuals whose post-mortem brain tissues carry histopathological signatures of AD, but showed no symptoms of cognitive decline or any other AD clinical features during their lifespan as "asymptomatic AD subjects". These histopathological signatures are similar by their location to mild cognitive impairment (MCI) and AD. Asymptomatic AD is related to over growing of cortical neurons located in anterior cingulate gyrus and CA1 hippocampus (48). These individual's post-mortem brain tissues are also used in studies as high-pathology controls (hpC) (49,50).

2.1.2. Parkinson's disease

Parkinsonism is a result of motor region of stratum's decreased dopaminergic transmission, which is characterized by bradykinesia with an additional motor skill deficiency. In addition to motor skill deficiencies, other symptoms of non-motor skills are also seen. Non-motor skills including, cognitive impairment like hallucination and dementia, mood changes, sleep disorder, autonomic dysfunction, pain and sensory symptoms follow an increasing pattern throughout the disease and are a great factor on the disability and nursing home placement of the patient. Instead of motor skill deficiencies, sleep disorder, depression constipation, anxiety and hyposmia are early signals of PD (Figure 2.2) (3).

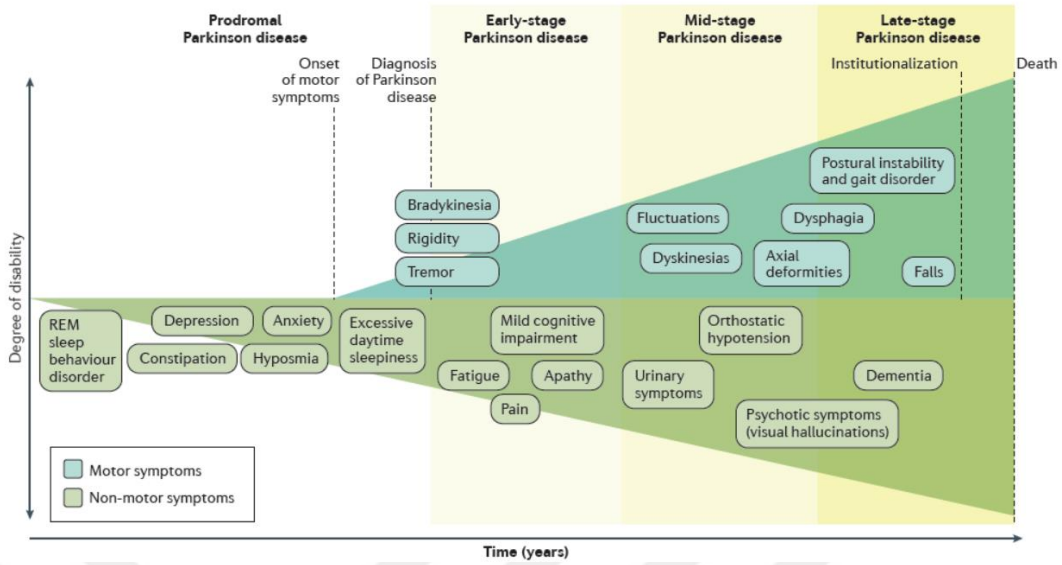

Figure 2.2. Motor and non-motor deficiencies of PD. The distribution of motor symptoms (blue) and non-motor symptoms (green) are summarized according to the disease stages (3).

The motor and non-motor deficiencies and their unparalleled development through the disease progression requires a complex staging manner. Usually, three staging scales are used in the same time for a more precise staging. The progress of the symptoms and disability level are measured by Hoehn and Yahr Scale. The scale was initially divided into five but since then intermediate stages were introduced with the better understanding of the disease (51). Unified PD Rating Scale (UPDRS) is a systematic scaling divided into 4 categories which are then divided into subsections; motor (part I) and non-motor (part II) aspects of experiences of daily living, motor examination (part III) and complications (part IV) (52). Daily activity preformation of the patients is measured by "Schwab and England Activities of Daily Living (ADL) Scale" (53). Scorings of these scaling systems were given in Table 2.1.

Table 2.1. Scoring of PD staging models. Hoehn and Yahr Scale uses stages from 0 to 5 while UPDRS uses levels from normal to severe. ADL system uses a percentage distribution system.

Hoehn and Yahr Scale	UPDRS	ADL
Stage 0 - No signs of disease Stage 1 – Unilateral symptoms Stage 1.5 - Symptoms unilateral and also involving the neck and spine Stage 2 - Symptoms on both sides but no impairment of balance Stage 2.5 - Mild symptoms on both sides, with recovery when the 'pull' test is given (the doctor stands behind the person and asks them to maintain their balance when pulled backwards) Stage 3 - Balance impairment, mild to moderate disease, physically independent Stage 4 - Severe disability, but still able to walk or stand unassisted Stage 5 - Needing a wheelchair or bedridden unless assisted.	For each subsection; 0: Normal 1: Slight 2: Mild 3: Moderate 4: Severe	100% = Completely independent. 90% = Completely independent but duration of activity increases. 80% = Mostly independent. 70% = Not completely independent. More difficulty with some chores. Three to four times as long in some. Must spend a large part of the day with chores. 60% = Some dependency. 50% = More dependent. 40% = Heavily dependent. 30% = Mostly dependent. 20% = Severely dependent. 10% = Completely dependent. 0% = Vegetative state.

Widespread intracellular α-synuclein accumulation and substantia nigra region's specific areas neuronal loss are the two most distinct histopathological signatures of PD. These signatures, when they appear alone, are not an indicator of PD, their coexistence is the hallmark of the disease. Lewy bodies, made of aggregated α-synuclein proteins are seen in cholinergic, monoaminergic and olfactory system neurons and also in neocortical and limbic brain regions. If the PD patient also has AD pathology the Lewy bodies mostly concentrate in limbic brain regions. Differing from Lewy body localization only in particular brain regions neurodegeneration is seen. Ventrolateral substantia nigra is the first effected region in this disease. The dopaminergic neurons are lost in this region while the neurons in other regions are protected. Parallel to the disease progression the loss of the neurons increase (3).

Ethnicity and gender are great risk factors for PD. When compared, Europe, North America and South America showed more tendency of having PD then Asia, Africa and Arabian Peninsula. Men with PD is 1.5 times more than women with PD and the progression of disease increases exponentially with age especially after 80 (4). Environmental factors contributes to the disease both as increasing and decreasing factors. Traumatic head injury, β-blocker usage, pesticide exposure and agricultural occupation are increasing factors while alcohol consumption, coffee drinking, tobacco smoking, calcium channel blocker usage are considered as decreasing factors (54). Genetic causes of the familial PD are divided into two groups according to the inheritance types; dominant and recessive. The familial PD is only build 5% of the PD disease but it lights the understanding of the disease. The frequencies of these gene mutations differ from very rare to common. Leucine rich repeat kinase 2 (*LRRK2),* glucosylceramidase beta *(GBA)* (common), DNA polymerase gamma (*Polγ)* (Rare), ubiquitin carboxy-terminal hydrolase L1 (*UCHL1),* HtrA serine peptidase 2 *(HTRA2),* eukaryotic translation initiation factor 4 gamma 1 *(EIF4G1),* DnaJ heat shock protein family (Hsp40) member C13 *(DNAJC13),* transmembrane protein 230 *(TMEM230)* (unclear) and synuclein alpha (*SNCA),* VPS35 Retromer Complex Component *(VPS35)* (very rare) shows autosomal dominant inheritance. Genes with autosomal recessive inheritance are; PTEN-induced kinase 1 (*PINK1),* Parkin RBR E3 Ubiquitin Protein Ligase *(PRKN),* Phospholipase A2 Group VI *(PLA2G6),* Vacuolar Protein

Sorting 13 Homolog C *(VPS13C)* (rare) and ATPase Cation Transporting 13A2 (*ATP13A2),* DnaJ Heat Shock Protein Family (Hsp40) Member C6 *(DNAJC6),* Parkinsonism associated deglycase *(PARK7), F-box protein 7 (FBXO7),* Synaptojanin 1 *(SYNJ1)* (very rare) (55). Some of the genetic causes of familial PD are linked to sporadic PD as risk factors. Although the mechanism underlying the sporadic PD has not been completely understood yet, the genetic factors contribution is a highly studied area. As a result of GWAS studies, *SNCA, LRRK2, GBA,* and *VPS13C* genes which are related to familial PD also showed significant relation with sporadic PD (55).

2.2. DNA Damage and Base Excision Repair in Neurons

Considering the lack of self-replicating ability of neurons and their life spans, DNA repair mechanisms play an important role of protecting the genome integrity and preventing death of neurons. DNA repair ensures that the neuronal operations function properly throughout life. Brain is protected by the blood-brain barrier and cranium from the exogenous DNA damaging factors. However these protections do not protect brain against spontaneous decays and reactions caused by cellular respiratory products, reactive oxygen species (ROS). Base excision repair (BER) is responsible from oxidative base damages, base modifications and single strand breaks. Because of the high metabolism rate, oxidative stress affects brain more than other organs and thus BER plays primary role in protecting integrity of brain cells (56).

BER pathway functions for both nuclear and mitochondrial DNA (mtDNA) and both are highly similar in operation. Different enzymes take place according to the location of the repair. BER mechanism like other DNA repair mechanisms starts with the identification and excision of the DNA damage. DNA glycosylase enzymes play a role in this first step and each enzyme is specific to the DNA damage but they frequently share overlapping specificity. For instance, 8-oxoguanine DNA glycosylase (OGG1) and 8-hydroxyguanine (8-OHGua), endonuclease VIII-like DNA glycosylase 1 (NEIL1) and 2,6-diamino-4-hydroxy-5-formamidopyrimidine (FapyGua), 4,6-diamino-5-formamidopyrimidine (FapyAde) and 5-hydroxyuracil (5-OHU), and uracil DNA glycosylase (UNG) for uracil. DNA glycosylase enzymes recognize the damaged base and cleave the N- glycosidic bond between the 2′-deoxyribose and base for

excising and thus creating the abasic site. The DNA glycosylases are divided into two groups: mono-functional and bi-functional. Mono-functional DNA glycosylases only excise the base and create abasic site. In addition to excision, bi-functional DNA glycosylases create a SSB with a non-conventional 3'-terminus. β-elimination or β,δ-elimination removes the non-conventional 3'-terminus. The elimination type determines the second step enzymes. Mainly AP endonuclease 1 (APE1) incises the backbone and creates strand break. This strand break has a non-conventional 5'-deoxyribose phosphate (dRP) and a priming 3′-OH group. For β,δ-elimination bi-functional glycosylases polynucleotide kinase-phosphatase (PNKP) incises the backbone and after incision a 3′-OH and 5′-PO4 end are generated. BER can take role as short-patch BER (SP-BER) and long-patch BER (LP-BER) from the repair synthesis step. DNA polymerases synthesize the nucleotide for filling the gap and also responsible from the dRP lyase activity after APE1 incision. For SP-BER DNA polymerase β (POLβ) is the main polymerase for short nucleotide gaps in nucleus and DNA polymerase γ (POLγ) in mitochondria. After the gap filling activity DNA ligase IIIα (LIGIIIα) X-Ray repair cross complementing 1 (XRCC1) complex completes SP-BER pathway by sealing the nick. For LP-BER DNA polymerase delta (POLδ) and DNA polymerase epsilon (POLε) perform strand displacement for synthesizing the base in conjunction with proliferating cell nuclear antigen (PCNA). After strand displacement, for creating a ligatable substrate the strand is excised by flap endonuclease 1 (FEN1). In LP-BER DNA ligase I (LIG1) seals the nick for completing the mechanism (57,58). BER pathway is represented in Figure 2.3.

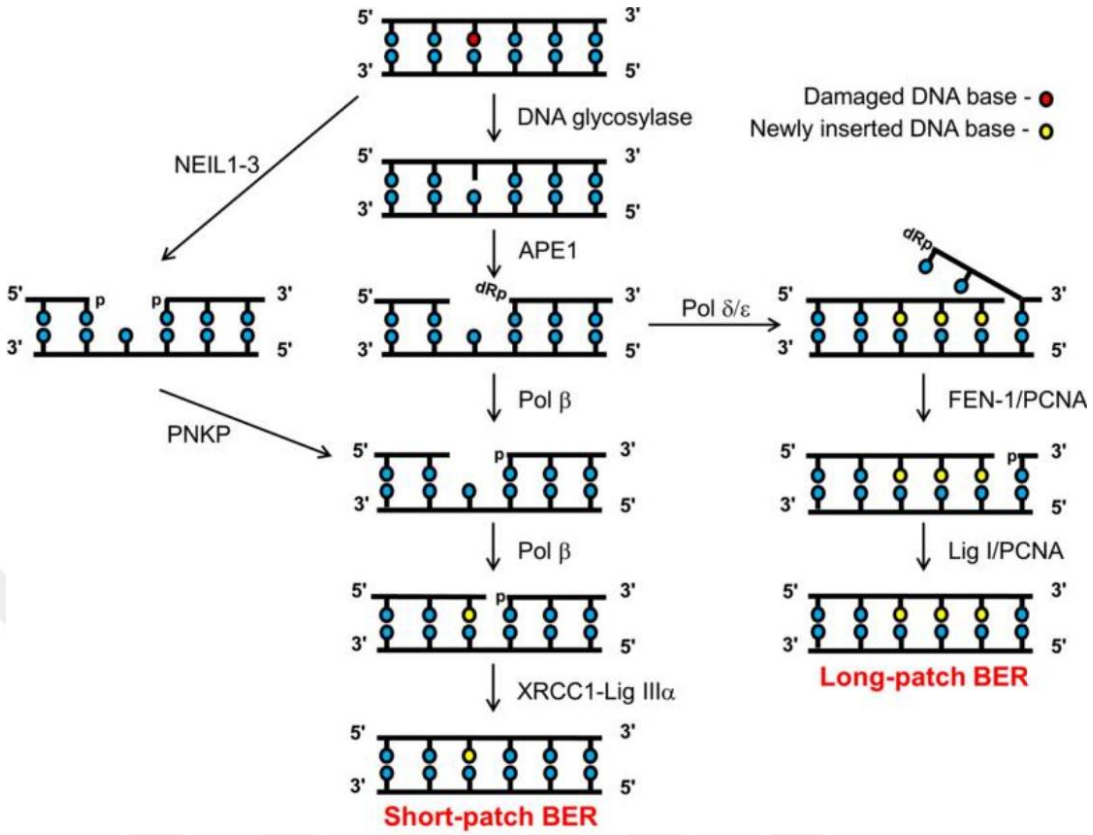

Figure 2.3. Base excision repair pathway. The steps and differences between both mono and bi functional polymerases and SP, LP BER are summarized (59).

The underlying reason of the switch between SP and LP BER has not been completely solved yet. The effect of the ATP concentration after dRP removal step was linked with this switch. In high ATP concentrations the pathway is completed by LIGIIIα, thus the SP-BER is selected by the cell (60). The protein-protein interactions, initiating lesion, differentiation status of the cell and POLβ resistant lesions like C1'-oxidized abasic lesion 2-deoxyribonolactone are shown to be a switching trigger between SP and LP BER (61,62).

2.2.1. Late-Onset Alzheimer's disease and base excision repair

Studies examining the relationship of oxidative DNA damage with LOAD, increase in oxidative DNA damage of blood samples and post-mortem brain tissues taken from individuals with LOAD was performed (5-7). Comparison of oxidized purine and pyrimidines in LOAD patient's and control's peripheral blood samples, has revealed a significant increase in LOAD samples. Likewise, oxidized purine and pyrimidine levels were found increasing in the nuclear DNA of frontal, temporal,

parietal lobes and cerebellum of post-mortem tissue samples (5,7). The DNA and RNA 8-OHGua levels and also OGG1 levels in post-mortem tissues of pre-clinical LOAD patients were higher in comparison to the control individuals (7). The incision activity of UNG and also its protein levels were found significantly lower in LOAD patient's post-mortem brain tissues in comparison to the control samples. The gap filling activity and POLβ protein levels were also lower in LOAD in the same samples. Consequently the total BER capacity was found significantly lower in LOAD patients (5). The 8-OHdG levels were higher LOAD patients nuclear DNA of spinal fluids when compared with controls. In the same samples free 8-OHdG was also higher which indicates oxidative stress yet alone is not a mark for LOAD but also BER mechanism deficiency is also a contributing factor (8). These studies using blood and tissue samples of both early and late stage LOAD patients, indicate the significant relation between early stages of LOAD pathogenesis and oxidative stress even in early stages of disease.

For further understanding of the relation between BER and LOAD, BER enzymes were studied. The APE1 knockdown sensory neurons were treated with platinum adducts increasing factor, cisplatin. Normally, nucleotide excision repair (NER) is responsible from this damage but in the APE1 knockdown cells, reduction of APE1 cells altered the RPA70 and XPA NER enzymes level (63). Glutamate signalization, which is a significant central nervous system neurotransmitter, is related with increased APE1 expression. Until a certain level it serves as a protecting factor against diseases for neurons but if the exposure duration prolongs, it can be toxic for neurons (64). PARP1 and APE1 inhibition in primary hippocampal neurons of mouse embryos showed that the BER mechanism may play a downstream role of DNA regulation for regulating synaptic transmission (65). Poly-(ADP-ribose) polymerase 1 (PARP1) meditates the PARylation which initiates the DNA damage response. The increased levels of PARP1 are related to cell death, because it reduces NAD+ levels. In LOAD patients post-mortem brain tissues PARP1 levels were found to be increased because of Aβ's, PARP1 enhancing factor. Also the PARylation level in post-mortem brain tissues was higher according to controls (66). Gene expression of *OGG1* in LOAD post-mortem brain tissues and MCI was significantly lower, when compared with control group. The decreased level of OGG1 was related to NFTs and dystrophic

neuritis. The decreased levels of MCI indicate the OGG1 relation with the early stage of LOAD (67). Also another study on post-mortem brain tissues showed that total BER activity was decreased in LOAD patient's samples in a correlation with UNG and POLβ activities, and also OGG1 activity was found decreased (5). Soltys and colleagues showed UDG activity in temporal cortex post-mortem tissues of LOAD patients were decreased in both nucleus and mitochondria and only at nucleus in cerebellum. The activity level of UNG showed no change in hpCs. This result indicates that UDG activity has an impact on disease outcome (50). In post-mortem inferior parietal tissues of LOAD patients, NEIL1 protein level and activity were found significantly lower when compared with control samples and due to decreased activity, 5-OHU levels were found higher in the same samples. Also LIGIIIα levels were significantly lower in mitochondrial extracts of the same samples (17). In LOAD patients the mRNA levels of BER genes *APE1, OGG1, MUTYH, PARP1* and *NEIL1* were in significant down regulation according to controls (69). Sykora and colleagues have developed a 3xTg/Polβ$^{(+/-)}$ mice strain to study the relation of POLβ and LOAD. The mouse was generated by crossbreeding 3xTgAD mice, a widely used LOAD mouse model and Polβ$^{(+/-)}$ mice. The reduced levels of POLβ in 3xTg/Polβ$^{(+/-)}$ mice resulted with DNA strand breaks and apoptosis elevation. The mice showed deficiencies in memory development and hippocampal synaptic plasticity. Autophagy deficiency increased levels of Aβ accumulation. Also the comparative transcriptome analysis of 3xTg/Polβ$^{(+/-)}$, 3xTgAD and Polβ$^{+/-}$ against human LOAD patients 3xTg/Polβ$^{(+/-)}$ results showed remarkable similarity to human LOAD patients (21). In cerebellum samples of LOAD patients, POLβ mRNA and protein were significantle elevated (70). Using neuron cell culture, the importance of POLβ in early stages of LOAD as a protective factor against Aβ was shown. In the same study a relation between Braak stages and POLβ expression was revealed. When Braak stage V-VI patients expression levels were found significantly low compared with Braak stage III/IV patients levels (71). The BER enzyme and LOAD phenotype relations were summarized in Table 2.2.

Table 2.2. BER and LOAD phenotype relation summary. The relation between BER enzymes levels, their effects on DNA repair and phenotype related to LOAD was summarized.

Gene	Phenotype	DNA repair and LOAD relation	Ref.

APE1	Glutamate signalization	↑DNA repair	↓Glutamate	↓LOAD	64
	Synaptic function deficiency	↑DNA repair	↑Synaptic regulation	↓LOAD	65
PARP1	Synaptic regulation deficiency	↑DNA repair	↑ Synaptic regulation	↓LOAD	65
	Neuronal loss	↑Aβ	↑PARP1	↑LOAD	66
OGG1	NFT	↓DNA repair	↑Hyperphosphorylated TAU	↑LOAD	67
UDG	Disease outcome	↓*UDG*	↓DNA repair	↑LOAD	50
NEIL1	Disease outcome	↓NEIL1	↓DNA repair	↑LOAD	17
POLβ	Aβ accumulation and plaque formation	↓DNA repair	↓↑Aβ	↑LOAD	21
	Neuron loss	↑DNA damage	↓Aβ	↓LOAD	
LIGIIIα	Disease outcome	↓ LIGIIIα	↓DNA repair	↑LOAD	17
↑ Increase, ↓ Decrease					

Some BER enzyme gene polymorphisms were studied to reveal the relation between genotype and phenotype. But the studies gave conflicting results on the risk factors of BER gene variations in LOAD. In a 105 LOAD 130 control cohort *APE1* rs1760944 T/T variant and rs1130409 Asp148Glu G/T and T/T variants were found significantly related to LOAD risk. The protective effect of rs1760944 T/G variant was shown in LOAD in a group of patients and controls with increased DNA damage. (72). *APE1* rs1760944 T/G variation was studied by the same group in another cohort with increased numbers of case and controls and no significant relation was found alone. But rs1760944 G/G in association with *LIGIII* rs1052536 C/T was significantly related to LOAD risk ($p<0.05$). In the same study rs1130409 gene variation G allele was also significantly related increased risk of LOAD while T allele was related decreased risk of LOAD (73). In a Turkish study population no significant relation was found between *APE1* rs1130409 and AD (74). The T allele of *APE1* rs1130409 was related with cognitive function decline in LOAD patients. MMSE score < 24 patients more likely to have T/T.genotype rather than G/T and G/G genotypes (70). *OGG1* rs1052133 gene variant was also studied in LOAD. Like *APE1* rs1760944, *OGG1* rs1052133 C/G variation was found significantly related to LOAD in a smaller study group but when the number of participants was increased, the significance of rs1052133 variation was lost. But the gene-gene interactions of *OGG1* rs1052133 and *NEIL1* rs4462560, *LIG3* rs4796030 and *LIG3* rs1052536 were found significantly related to LOAD. Also the

rs1052133 gene variation was significantly related to DNA damage repair decrease (72,73). Other studies, using PCR-RFLP for rs1052133 also couldn't find a significant relation between *OGG1* and LOAD (74,75). But Dorszewska and colleagues was found a significant relation between *OGG1* rs1052133 and increased 8-oxodG levels in LOAD. No significant relation between *OGG1* rs104893751 and LOAD was found (76). In LOAD post-mortem brain tissues, *OGG1* rs3219012 (Ala288Val) and Ala53Thr polymorphisms were related to reduced OGG1 glycosylase activity (77). *NEIL1, FEN1, LIGI, LIGIII, XRCC1, PARP* and *MUTYH* gene variations were also studied by Kwiatkowski and colleagues. *NEIL1* rs4462560 C/G, *FEN1* rs174538 G/G and *LIGI* rs20579 G/G variations were significantly related to LOAD when studied in a small cohort (105 case, 130 control). *LIGIII* rs4796030 C/C was related to decrease of DNA repair efficacy. In the larger cohort (281 case, 150 control), no significant relation was found between *NEIL1* and *FEN1* gene variations. *LIGI* rs20579 A/A, *LIGIII* rs1052536 C/C, rs4796030 A/C variations were related to increased risk while *LIGI* rs20579 G/- and *LIGIII* rs4796030 C/C was related to decreased risk of LOAD (72,73). *XRCC1* rs25489 and rs25487 gene variations showed no significant association in LOAD patients in a 91 case 93 control study group (74). In Kwiatkowski and colleagues 2015 study, *XRCC1* rs25487 G/A variation was related to increased risk a and A/A variation was related to decreased risk of LOAD. They couldn't find a significant relation between rs1799782 and LOAD in 2015. But in 2016 study, rs1799782 C/T was related to increased risk and DNA damage in LOAD while T/T showed a protecting factor against LOAD. rs25487 G/A was also found as an increasing risk factor of LOAD and related to increased DNA damage in LOAD (72,73). The *XRCC1* rs1799782 gene variation was also studied by two other group using PCR-RFLP method but no significant association was found between the variation and LOAD (79,80). *PARP1* gene rs1805404 and rs1136410 variations showed no significant association alone but when haplotype (Ht) analysis was performed between rs1805404 C/T and rs1136410 T/C, Ht1 (TC) was found as a decreasing factor while Ht3 (TT) was found as an increasing factor of LOAD (81). On the contrary, rs1136410 T/C gene variation was found as an increasing risk factor for LOAD while T/T was found as decreasing factor. *MUTYH* gene rs3219489 variation was studied in two different cohorts and no significant relation was found for LOAD (72,78). The relation of BER gene variations and LOAD was summarized in Table 2.3.

Table 2.3. BER and LOAD genotype relation summary. The relation between BER gene variants and LOAD, their amino acid change, significance and case/control numbers, the methodology and sample type was summarized.

Gene	Pathway	rs code	Polymorphism	Amino Acid Change	Significance	Case /Control	Method	Sample Type	Reference
APE1	BER	rs1760944	c.-468T>G	N/A	T/T ↑ T/G ↓	105/130	SNP Genotyping	Peripheral Blood	72
		rs1130409	c.444T>G	Asp148Glu	T/G ↑ T/T ↑				
		rs1760944	c.-468T>G	N/A	-	281/150	SNP Genotyping	Peripheral Blood	73
		rs1130409	c.444T>G	Asp148Glu	G allele ↑ T allele ↓				
		rs1130409	c.444T>G	Asp148Glu	-	91/93	PCR-CTPP	Peripheral Blood	74
		rs1130409	c.444T>G	Asp148Glu	Reduced cognitive performance	161/96	SNP Genotyping	Peripheral Blood	70
OGG1	BER	rs1052133	c.977C>G	Ser326Cys	DNA damage↑	105/130	SNP Genotyping	Peripheral Blood	72
		rs1052133	c.977C>G	Ser326Cys	-	281/150	RT-PCR	Peripheral Blood	73
		rs1052133	c.977C>G	Ser326Cys	-	91/93	PCR-RFLP	Peripheral Blood	74
		rs1052133	c.977C>G	Ser326Cys	-	178/146	PCR-RFLP	Peripheral Blood	75
		rs1052133	c.977C>G	Ser326Cys	C/G ↑	41/51	PCR-RFLP	Peripheral Blood	76
		rs104893751	c.137G>A	Arg46Gln	-	41/51	PCR-RFLP	Peripheral Blood	
		-	-	Ala53Thr	Associated with reduced OGG1 activity	14/10	SSCP and DNA sequencing	Post-mortem brain tissue	77
		rs3219012	c.863C>T	Ala288Val					
NEIL1	BER	rs4462560	c.*283C>G	N/A	C/G ↑ -	105/130 281/150	SNP Genotyping	Peripheral Blood	72 73
FEN1	BER	rs174538	c.-441G>A	N/A	G/G ↑ -	105/130 281/150	SNP Genotyping	Peripheral Blood	72 73

Gene	Pathway	rs code	Polymorphism	Amino Acid Change	Significance	Case /Control	Method	Sample Type	Reference
LIGI	BER	rs20579	c.-7C>T	N/A	G/G↑ A/A↑ G/- ↓	105/130 281/150	SNP Genotyping	Peripheral Blood	72 73
LIGIII	BER	rs1052536	c.*50C>T	N/A	-				
		rs4796030	c.*83A>C	N/A	A/C ↑ C/C ↓	281/150	SNP Genotyping	Peripheral Blood	73
					-	105/130			72
XRCC1	BER	rs25489	c.839G>C	Arg280His	-	91/93	PCR-RFLP	Peripheral Blood	74
		rs25487	c.1196A>G	Gln399Arg	-		Capillary PCR	Peripheral Blood	
		rs25487	c.1196A>G	Gln399Arg	G/A↑ A/A ↓	120/110	SNP Genotyping	Peripheral Blood	78
		rs1799782	c.580C>T	Arg194Trp	-				
		rs25487	c.1196A>G	Gln399Arg	G/A ↑ DNA damage	105/130	SNP Genotyping	Peripheral Blood	
		rs1799782	c.580C>T	Arg194Trp	C/T ↑ T/T ↓ DNA damage	105/130	SNP Genotyping	Peripheral Blood	72
		rs1799782	c.580C>T	Arg194Trp	-	98/95	PCR-RFLP	Peripheral Blood	79
		rs1799782	c.580C>T	Arg194Trp	-	212/203	PCR-RFLP	Peripheral Blood	80
PARP1	BER	rs1805404	414C/T	Asp81Asp	-	120/111	PCR-RFLP	Peripheral Blood	81
		rs1136410	2285T>C	Val762Ala					
		rs1136410	2285T>C	Val762Ala	T/C↑ T/T ↓	120/110	SNP Genotyping	Peripheral Blood	78
		rs1136410	2285T>C	Val762Ala	T/C↑ T/T ↓	105/130	SNP Genotyping	Peripheral Blood	72

Gene	Pathway	rs code	Polymorphism	Amino Acid Change	Significance	Case /Control	Method	Sample Type	Reference
MUTYH	BER	rs3219489	c.972G>C	Gln324His	- -	120/110 105/130	SNP Genotyping	Peripheral Blood	78 79
- No effect on LOAD risk, ↑ Increase LOAD risk, ↓ Decrease LOAD risk									

2.2.2. Sporadic Parkinson's disease and base excision repair

Oxidative damage was found severely decreased in substantia nigra neurons of PD patients using 8-OHGua immunoreactivity. This damage was especially linked with cytoplasmic nucleic acids, RNA and mtDNA (9). In sporadic PD patients and controls, several brain regions were studied to define the difference of oxidative DNA damage which were; cerebellum, frontal pole, substantia nigra, caudate nucleus, anterior putamen, medial putamen, posterior putamen, nucleus accumbens, substantia innominate and lateral globus pallid. In all different regions, FapyGua was decreased and 8-OHGua was increased. Especially in substantia nigra, these differences were significantly important (10). One of the hallmarks of PD, accumulated α-synuclein was linked to increased ROS in PD brain. Using electron microscopy, localization of α-synuclein in mitochondria was identified in mainly inner membrane of mitochondria. And the accumulation in mitochondria was found significantly higher in substantia nigra and striatum of PD patients compared to controls. Also, mitochondria complex I activity was showed 3 fold decrease in significant negative correlation with accumulated α-synuclein in the PD patient's same samples. Resulting of the accumulation a significant increase in reactive oxygen species was also detected (11). Oxidative stress inducer pesticide, paraquat was linked to PD pathogenesis. The increased oxidative stress in PD patients due to paraquat was related with redox cycling and also with α-synuclein accumulation (12).

The regulating role of Parkin, a PD-related protein on APE1 was identified. Using cell culture and glioblastoma tissue samples from PD patients a reverse correlation between Parkin and APE1 expression was shown (82). The mtDNA repairing enzyme hOGG1-2a levels were found significantly upregulated in substantia nigra region of PD patients post-mortem brain tissues when compared with progressive supranuclear palsy patients, corticobasal degeneration patients and also control subjects (83). A feed-forward loop was found between PARP1 enzyme and α-synuclein. It was shown that while α-synuclein as increasing the PARP1 level and thus PARylation, PAR generation increases the α-synuclein formation (84). Xue and colleagues studied on PD mice model generated by 6-hydroxydopamine (6-OHDA) and 1-methyl-4-phenyl-1,2,3,6-

tetrahydropyridine (MPTP) induced C57BL/6 *NEIL1* knockout (-/-) mice which are used for selectively targeting dopaminergic neurons. *NEIL1*(-/-) revealed significantly increased motor dysfunction and dopaminergic neuron loss. A positive relation was found between the survival of dopaminergic neurons and NEIL1 expression. Later in MPTP induced mice, inflammatory response was also found increasing and when NEIL1 expression promoted the PD progression was reversed (85). In the rotenone induced dopaminergic neuron cells and PD rat model DNA *POLβ* was overexpressed in substantia nigra of the PD model rats. Further the silencing of *POLβ* decreased the endoreduplication resulted from the rotenone treatment. These results suggest the role of *POLβ* cell cycle regulation may effect PD progression (22). The BER enzymes relation with sporadic PD was summarized in Table 2.4.

Table 2.4. BER and PD phenotype relation summary. The relation between BER enzymes levels, their effects on DNA repair and phenotyope related to PD was summarized.

Gene	Phenotype	DNA repair and PD relation			Ref.
APE1	PD phenotype	↑Parkin	↓ APE1	↑ PD	82
PARP1	PD phenotype	↑ α-synuclein	↑ PARP1	↑ PD	84
OGG1	PD phenotype	↑ oxidative stress	↑ *OGG1*	↑ PD	83
NEIL1	PD progression	↓NEIL1	↑ neuron loss	↑ PD	85
POLβ	PD progression	↑ *POLβ*	↑ endoreduplication	↑ PD	22
↑ Increase PD risk, ↓ Decrease PD risk					

APE1 rs1130409 and *OGG1* rs1052133 well known gene variations were studied in a large sporadic PD group for investigating their effects on PD pathogenesis. Both variations showed no significant association with PD by individually or in relation with each other. But in the presence of pesticide exposure, carrying these genetic variations was significantly contributed PD pathogenesis (12). *OGG1* rs1052133 variation was studied in a smaller group (139 PD and 211 control) and no significant association was also detected in PD group (86). In a Turkish cohort, BER gene variations was also studied, *APE1* rs1130409 G/G and *XRCC1* rs25487 T/T was significantly related to PD while *OGG1* rs1052133 showed no significant association (87). On the contrary in another study in Turkish cohort *XRCC1* rs25487 showed no significant association with PD patients (88). In a large (600 case 592 control) study group, *PARP1* variations selected according to the results of haplotype analysis were genotyped. The genotyped variations are; rs12567614, rs8679, rs3219142, rs3219123, rs3219110, rs1805410,

rs1805414, rs3219053, rs1805405, rs907187 and rs1136410. No significant association was found between *PARP1* gene variants and PD (89). POLγ has been related with familial PD previously. But in a Finnish population based study, POLγ non-common (other than 10 and 11Q long PolyQ) polyglutamine (PolyQ) track was significantly related with sporadic PD when compared with age matched spouses, population controls and non-parkinsonian neurodegenerative disorders (90). The relation of BER gene variations and PD was summarized in Table 2.5.

Table 2.5. BER and PD genotype relation summary. The relation between BER gene variants and PD, their amino acid change, significance and case/control numbers, the methodology and sample type was summarized.

Gene	rs Code	Polymorphism	Amino Acid Change	Significance	Case /Control	Method	Sample Type	Reference
APE1	rs1130409	c.444T>G	Asp148Glu	-	619/854	SNP Genotyping	Peripheral Blood	12
	rs1130409	c.444T>G	Asp148Glu	G/G ↑	60/108	PCR-RFLP	Peripheral Blood	87
OGG1	rs1052133	c.977C>G	Ser326Cys	-	619/854	SNP Genotyping	Peripheral Blood	12
	rs1052133	c.977C>G	Ser326Cys	-	139/211	PCR-RFLP	Peripheral Blood	86
	rs1052133	c.977C>G	Ser326Cys	-	60/108	PCR-RFLP	Peripheral Blood	87
XRCC1	rs25487	c.1196A>G	Gln399Arg	–	71/76	PCR-RFLP	Peripheral Blood	88
	rs25487	c.1196A>G	Gln399Arg	T/T ↑	60/108	PCR-RFLP	Peripheral Blood	87
PARP1	rs12567614	g.226544420T>C	N/A	–	600/592	SNP Genotyping	Peripheral Blood	89
	rs8679	c.2285T>C	Val762Ala	–				
	rs3219142	g.226552068G>A	N/A	–				
	rs3219123	g.226555348G>A	N/A	–				
	rs3219110	g.226557878T>C	N/A	–				
	rs1805410	g.226568665T>C	N/A	–				
	rs1805414	g.226573364A>G	Ala284Ala	–				
	rs3219053	g.226574554C>T	N/A	–				
	rs1805405	g.226580021G>A	N/A	–				
	rs907187	g.226595647C>G	N/A	–				
	rs1136410	2285T>C	Val762Ala	–				

- No effect on PD risk, ↑ Increase PD risk, ↓ Decrease PD risk

2.3. Human DNA Polymerase Beta

DNA polymerases which are responsible from nucleic acid synthesis are divided into 7 families according to their sequence homologies and crystal structures (A, B, C, D, X, Y and RT). *POLβ*, is responsible from gap filling DNA synthesis activity and belongs to X family and is present in all mammalian species as a 39 kDa peptide with high level sequence conservation. This level of sequence conservation indicates the role of POLβ in mammalian survival. *POLβ* gene is composed of 50 to 233 bp long 14 exons and 96 bp to 6.5 kb long 13 introns. There are two domains on the POLβ protein structure: 31 kDa polymerase and 8 kDa lyase domains. Polymerase domain is also divided into three subdomains: (D) dsDNA binding domain, (C) catalytic domain responsible from nucleotidyl transferase and (N) base pair binding domain responsible from nascent base binding. 5'-dRP lyase activity and ssDNA binding is lyase domain's responsibility. Figure 2.4A represents the scheme of POLβ from primary structure to domains. The major secondary elements, α-helices and β-sheets, and helix-hairpin-helix motifs. The exon boundaries were aligned with corresponding subdomains. The organizations of domains and subdomains of POLβ was shown in Figure 2.4B (91,92).

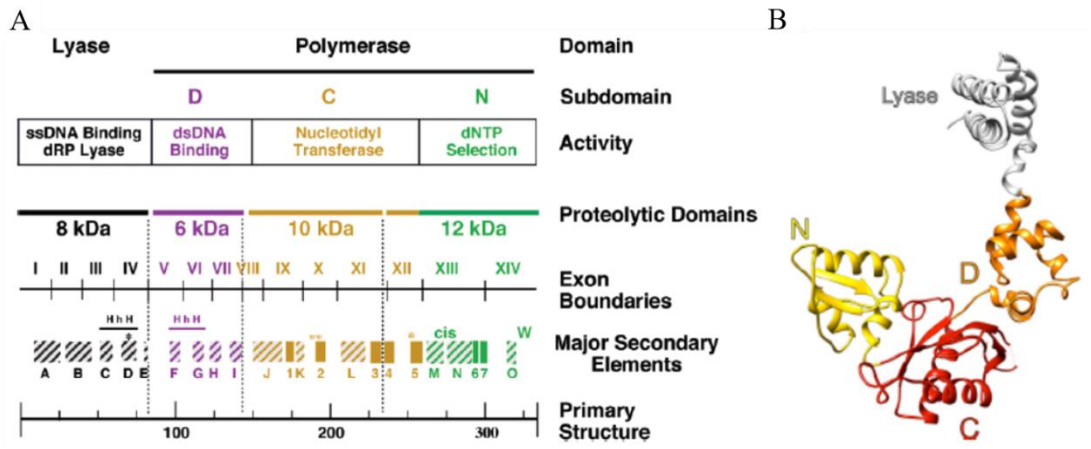

Figure 2.4. Structure of *POLβ* gene and protein. A. The scheme of POLβ from primary structure to domains. B. Organization of POLβ domains (91,92).

POLβ levels are found at higher levels in brains of mammals. Especially in embryonic period it is generally higher in all brain regions but in adulthood the high levels were localized in hypothalamus and cerebellum (93). Studies comparing

proliferating and non-proliferating cells showed that while polymerases δ and ε is decreasing in non-proliferating cells, POLβ levels were increased or remained stable. For further studying this decrease, LP-BER activity was compared using protein extracts of intestinal mucosa and post-mitotic dentate gyrus hippocampal regions. LP-BER activity was shown to be decreased in hippocampal tissue. Also POLβ was involved in an alternative PCNA independent LP-BER activity in the post-mitotic dentate gyrus POLβ (94). The results of these studies indicate the importance of POLβ in post-mitotic brain cells in both SP and LP BER.

POLβ (*POLβ* -/-) mice were used for investigating the role of *POLβ* in post-mitotic cells. In contrast with the studies using knock out core BER enzymes, *POLβ* -/- did not show embryonic lethality. The *POLβ* -/- mice were smaller according to other mice and they did not developed breathing reflex. This developmental death occurred in late stage was linked to neuronal differentiation problems. On E12.5 apoptosis was detected in developing nervous system and p53 activation was associated with apoptosis. This association suggested DNA damage is underlying under this apoptotic response. When the other organs of these mice was investigated no histological abnormality was detected which corroborates the relation between post-mitotic cells and *POLβ* (20). *POLβ* heterozygous (*POLβ* +/−) mice strain was developed and studied for *POLβ* and LOAD pathogenesis. The damage accumulation was found significantly higher in *POLβ* +/− and neuron deaths were found resulting from Aβ accumulation. Low levels of *POLβ* gene expression was also related to memory dysfunction and synaptic plasticity deficiency (21).

2.4. Human Uracil DNA Glycosylase

Uracil DNA glycosylase (UDG) family is responsible from incision of uracil in DNA, generated from cytosine deamination or misincorporation. Uracil in DNA is found to be potentially mutagenic and even small amounts may affect protein and DNA interactions. The activity of human UDG (UNG) was detected in all human tissues. Two forms of *UNG* was detected; mitochondrial UNG form, UNG1 and nuclear UNG form, UNG2. Both UNG forms were shown to be expressed from the same *UNG* gene (95). *UNG* gene has 89 to 1164 bp long 6 exons and 503 to 6449 long 5 introns. *UNG*

gene structure is shown in Figure 2.5. The localization of exons, restriction sites, interspersed repeats and also 300 bp TA repeat are represented (96).

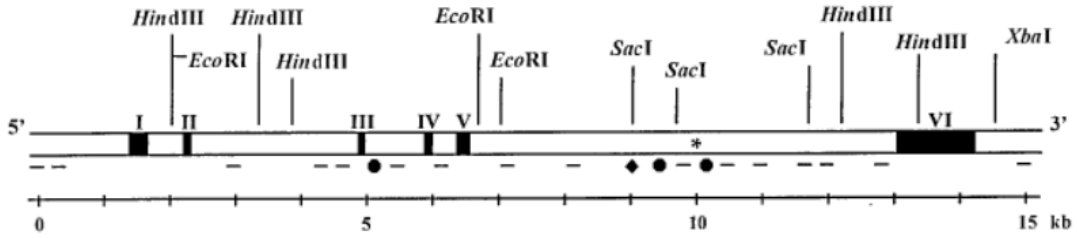

Figure 2.5. *UNG* gene structure. The exons are shown in black rectangular boxes. Alu: —, MER: hexagon boses, MIR: diamond boxes and * indicates 300 bp TA repeat (96).

UNG1 and UNG2 are alternatively spliced from the N-terminal region of *UNG*. EGFP was used for studying the difference in sequence of UNG1 and UNG2 N-terminal sequenced and in localizations. The comparison of UNG2 cDNA sequence with UNG sequence, revealed an unrecognized exon, which named exon IA. And the originally known exon I from first *UNG* sequencing was named exon IB. Both exons, IA and IB have different promoter sites; PA and PB, respectively. The N-terminal regions of UNG1 and UNG2 differ in size; 304 and 313 amino acids respectively. UNG1 has 35 amino acids different from UNG2 while, UNG2 has 44 amino acids different from UNG1. The splicing scheme was summarized in Figure 2.6. The splicing site in IB localizes after 35 amino acids from the beginning of IB. UNG1 is formed from the axon IB completely and remaining 5 exons. UNG2 is formed from exon IA and IB after splicing region (lack of 35 amino acids from the beginning). The localization sequence of UNG2 with bold letters is also shown in Figure 2.6 (97).

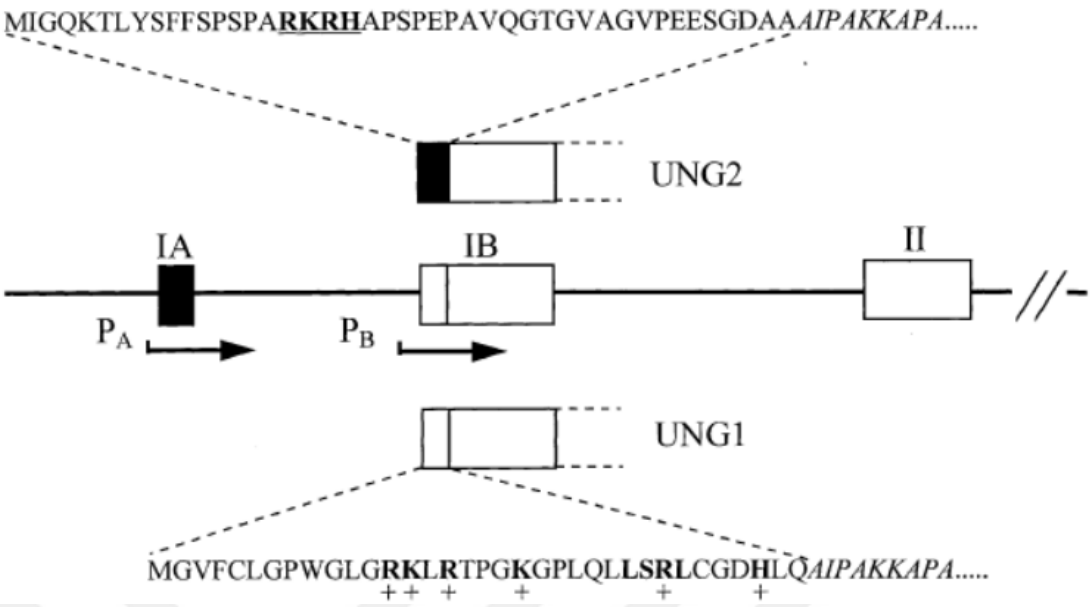

Figure 2.6. Alternative splicing of *UNG* gene. Exons IA and IB are shown on the gene structure. The combinations of UNG1 (exon IB) and UNG2 (exon IA plus IB) are also represented. The N-terminal and localization sequence of UNG2 (underlined) and N-terminal UNG1 sequence were showed (97).

Kruman and colleagues used folic acid deficiency as a triggering factor for uracil misincorporation and showed, the accumulation of uracil is related with neuronal cell death (13). When UNG activity was suppressed in rat embryonic hippocampal cells, the p53 activation, nuclear chromatin condensation and neurites fragmentation was significantly increased. And also the DNA damage, which is the underlying reason of this increased neuronal cell death was also showed (14). In a long term study of dementia free elderly, homocysteine which is related with deficient folic acid and cognitive decline was independently related. In the end of the study, 112 subjects developed dementia; including 70 LOAD was significantly related to hyperhomocysteinemia and low folic acid levels (15). Decreased levels of UDG in neuron cells and as a consequence reduced BER activity were showed to be related with increasing age which is one of the major factors of dementia.

2.5. Human Endonuclease VIII-like DNA Glycosylase 1

DNA glycosylases endonuclease VIII-like (NEI) mammalian family is constituted of; three Nei-like DNA glycosylases *NEIL1*, *NEIL2* and *NEIL3*. *NEIL1* is primarily responsible from removing oxidized pyrmidine bases such as; thymine glycol, formamidopyrimidine (Fapy) and 5-hydroxuracil 5-OHU. *NEIL1* gene has 28 to 456

bp long 11 exons and 93 to 1037 bp long 10 introns. The domains of *NEIL1* gene is shown in Figure 2.7. NEIL1 has four identified domains; conserved N-terminal, (Fpg/Nei superfamily), helix-two-turn helix (H2TH) Zinc finger and nuclear localization sequence (NLS). Zn domain is important for interacting with PCNA, LIGIII, POLβ and XRCC1 (98).

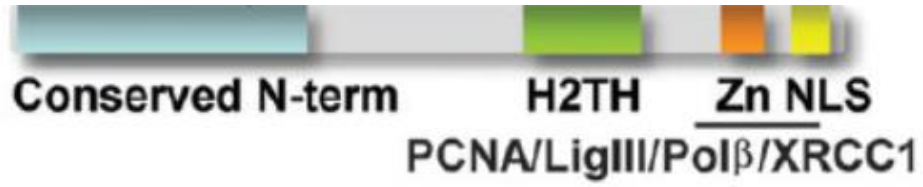

Figure 2.7. *NEIL1* gene structure. The domains of NEIL1 gene were shown on the structure. Blue: N-terminal, green: H2TH, orange: Znf and yellow: NLS (98).

FapyA one of the main targets of NEIL1 was found significantly elevated in liver, kidney and brains of when compared with control in old *neil1*$^{-/-}$ mice (16). *neil1*$^{-/-}$ mice was also used for studying the relation between NEIL1 and brain function. Using Morris water maze Canugovi and colleagues showed that their learning capacity was remained same with wild type (wt) mice but their memory retaining ability was decreased. They also used focal ischemic stroke on same strain and observed increased brain damage and mortality according to the wt mice. As a result of DNA damage comparison, after focal ischemic stroke, in *neil1*$^{-/-}$ mice DNA damage repair capacity was decreased when compared with wt mice. These results indicate the importance of NEIL1 enzyme activity in post-mitotic neuronal cell survival (68). Comparison of DNA glycosylase activities through-out the life span using mice revealed, rather than other glycosylases the expression of NEIL1 was increased with the age in hippocampus, cerebellum and neocortex (18). Later using the *neil1*$^{-/-}$ mice Canugovi and colleagues revealed the relation between *neil1*$^{-/-}$ deficiency and olfactory function which has been linked to the neurodegenerative disorders. *NEIL1* deficiency revealed decreased olfactory function among *neil1*$^{-/-}$ mice (19).

2.6. Human Apoliprotein E

APOE gene is one of the genes responsible from cholesterol and lipid metabolism and transport. APOE primarily function in liver, kidney and spleen and in central

nervous system it is found in microglia, astrocytes and in small portion neurons. In brain, *APOE* takes role in repair by lipid redistribution among neurons. *APOE* is composed of 66 to 861 bp long 4 exons and 580 to 1092 bp long 3 introns. The two domains; receptor binding and lipid binding are joined to each other with a flexible hinge region. Two domains interact with each other from Arg61 and Glu255 (Figure 2.8) (99).

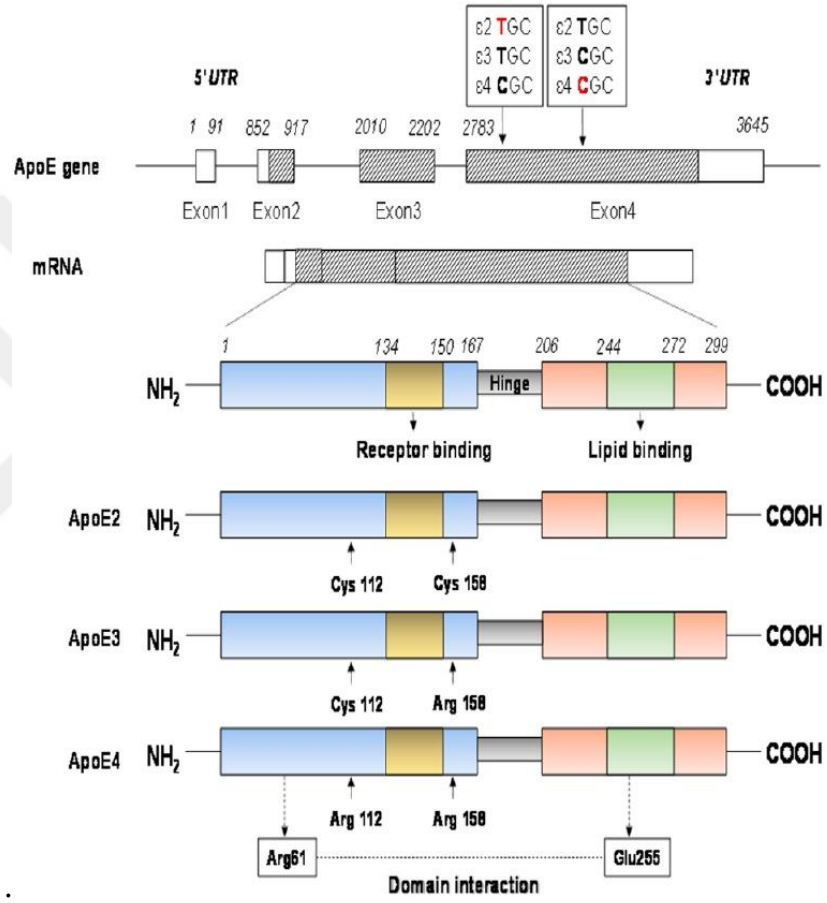

Figure 2.8. *APOE* gene structure. The exons, mRNA form and domains of APOE are represented in the figure. APOE2, APOE3 and APOE4 differences by 112 and 158 residues are also included (99).

Human *APOE* is found in 3 different isoforms; ε2, ε3 and ε4. The isoforms are determined according to the amino acid change at 112 and 158 residues. The rs429358 and rs7412 variations are linked to these isoforms. Table 2.6 summarizes the alleles and amino acids of APOE isoforms. The isoforms have different lipid and receptor binding affinities. The low density lipoprotein (LDL) binding affinity of ε2 is much lower according to the other isoforms. ε4 binds to triglyceride enriched proteins while

ε2 and ε3 preferentially bind phospholipid enriched LDLs (99). ε3 has the highest frequency (60–70%) in healthy humans according to the other alleles followed by ε4 (15–20%) and ε2 (5–10%). ε4 has been linked to LOAD pathology. In LOAD, ε4 frequency increases to 50% and in contradiction to ε4, ε2 has been linked to protective effect to *APOE* against LOAD. ε4 also linked to cognitive decline, with decreased Aβ clearance and greater brain atrophy, decreased glucose mechanism in cerebra, synaptic function impairment and hippocampal neurogenesis defect (38). The role of the *APOE* genotype; ε2/ε2, ε3/ε3, ε4/ε4, ε2/ε3, ε2/ε4, ε3/ε4 was also studied in LOAD. ε4 allele contributing factor in genotypes was also showed. With each ε4 allele risk AD was increased by 2.8 fold. So the ε4/ε4 genotype carriers have the greatest risk for AD (100). Genotype details of *APOE* are also summarized in Table 2.6. *APOE* was also related with sporadic PD. In PD patients both ε2 and ε4 was found significantly related and also both of them was found significantly related to PD related dementia (101,102).

Table 2.6. *APOE* alleles. ε2, ε3 and ε4 relation with rs429358 rs7412 gene variation was summarized on allele and amino acid level.

	rs429358		rs7412	
	Allele			
	Allele	Amino acid	Allele	Amino acid
ε2	T	Cys 112	T	Cys 158
ε3	T	Cys 112	C	Arg 158
ε4	C	Arg 112	C	Arg 158
	Genotype			
	Genotype	Amino acid		
ε2/ε2	TT/TT	Cys 112-Cys 158		
ε3/ε3	TC/CC	Cys 112-Arg 158		
ε4/ε4	CC/CC	Cys 112-Arg 158		
ε2/ε3	TT/TC	Cys 112-Cys 158-Arg 158		
ε2/ε4	TC/TC	Cys 112-Arg112-Cys 158-Arg 158		
ε3/ε4	TC/CC	Cys 112-Arg112-Arg 158		

2.7. Sequencing Technologies and Importance in Diagnosis of AD and PD

Sanger sequencing is the first generation of new generation sequencing technologies which uses chain-terminating dideoxynucleotides by DNA polymerase to reveal the sequence of amplified target gene regions. It is still the gold standard among sequencing technologies but with the development of parallel sequencing technologies,

next generation sequencing technologies (NGS) generate more rapid, robust and high-throughput results. The parallel sequencing gives chance to analyze multiple regions in multiple individuals in a relatively short time. And this time reducing effect and high-throughput results are one of the most important reasons for choosing the NGS for diagnostics (103).

Neurodegenerative disorders are difficult to clinical diagnosis due to their heterogenous and overlapping syndromes. In LOAD 24% and in sporadic PD over 25% of patients is misdiagnosed likewise other neurodegenerative disorders. These misdiagnosis rates lead to a search for new diagnostic tools and genetic markers for diagnosis, NGS serves a new promising and most importantly non-invasive option for neurodegenerative disorders diagnosis.

3. MATERIALS AND METHODS

3.1. Study Population and Sample Collection

The peripheral blood samples of 198 LOAD patients (≥65 years), 50 sporadic PD patients and age-matched 98 cognitively normal controls without any AD family history were recruited at the department of Neurology, Medeniyet University Goztepe Training and Research Hospital, Istanbul, Turkey. In addition, DNA samples from postmortem brain tissues (temporal cortex (TC) and cerebellum (CE)) of LOAD (11 CE, 10 TC), cognitively normal control without AD pathology (10 CE, 9TC) and high-pathology control (hpC, 11 CE, 8 TC) cognitively normal with high AD neuropathological changes) were kindly provided by Dr. Nadja Souza Pinto, University of São Paulo, Brazil (the Brazilian Aging Brain Study Group's Brain Bank, University of São Paulo, School of Medicine) Written informed consents were collected from all subjects prior to participation in this study. The study was approved by the Ethics Committee of Acibadem Mehmet Ali Aydinlar University and Acibadem Health Institutions Medical Research (B.30.2.ACÜ.0.00.00.050-06/45). The clinical diagnosis of LOAD was made according to the Neurological and Communicative Disorders and Stroke-Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) criteria and the criteria of Diagnostic and Statistical Manual of Mental disorders, 4th ed. (DSM-IV). Cognitively normal participants received the same assessment as the cases, and were accepted non-demented. Ten ml of blood taken from each participants were collected in the BD Vacutainer Blood Collection Tubes that contains EDTA (Becton, Dickinson and Company, USA). Blood samples were stored at 4°C until cold chain transfer. DNA isolation was performed at Acibadem Mehmet Ali Aydinlar University.

Blood samples of same six LOAD participants and four control participants were collected for gene expression analysis from the department of Neurology, Medeniyet University Goztepe Training and Research Hospital, Istanbul, Turkey. The patients and controls for gene expression analysis were selected according to the bioinformatics

analyses. Three ml of blood from participants were collected in the BD Vacutainer Blood Collection Tubes that contains EDTA (Becton, Dickinson and Company, USA). Blood samples were stored at 4°C until cold chain transfer. RNA isolation was performed at Acibadem Mehmet Ali Aydinlar University.

3.2. DNA, RNA Isolation and Quality Control

Total DNA isolation from blood samples was performed using QIAamp DNA Blood Kit (Qiagen, USA). Up to 200 µL of blood samples were added onto 20 µL of Proteinase K in a 1.5 mL microcentrifuge tube. Buffer AL (200 µL) was added onto samples, then mixed by pulse vortexing for 15 seconds and the mixtures were incubated at 56ºC for 10 minutes. After incubation, samples were centrifuged for short period of time and 200 µL of 100% ethanol was added into tube. The mixtures were transferred into QIAamp Mini spin column and centrifuged at 6000xg for 1 minute. The spin columns were placed into a new collection tube, 500 µL of Buffer AW1 was added into spin columns and centrifuged at 6000xg for 1 minute. Spin columns were placed into a new collection tube, 500 µL of Buffer AW2 was added into spin columns and centrifuged at 20,000xg for 3 minutes. After centrifugation, spin columns were placed into a new collection tube and centrifuged at full speed for 1 minute. Then, spin columns were placed into a 1.5 mL microcentrifuge tube and 200 µL dH2O was added into spin columns. Samples were incubated at room temperature for 3 minutes and centrifuged for DNA elution at 6000xg for 1 minute. The DNA samples were stored at -20°C until further usage.

Peripheral Blood Mononuclear Cells (PBMC) were isolated from blood samples for RNA isolation. In a 15 mL falcon, 6 ml Histopaque®-1077 (Sigma Aldrich, USA) was added. EDTA tubes were mixed by inverting 5 times and 3 ml blood samples were mixed with 3 ml PBS. The mixture was also mixed by inverting 5 times and all 6 mL was added onto 6 ml Histopaque by tilting the Histopaque falcon tube by 45 degrees and leaning the serological pipette against the falcon tube wall. The mixture was poured very slowly and carefully for not breaking the liquid wall. After pouring was completed, the mixture was centrifuged at 400xg for 30 minutes with acceleration 9

and deceleration 1. The ~1.5 mL mononuclear cell layer was transferred to a clean 15 mL falcon tube and mixed with 6 ml PBS by inverting 5 times for washing. The mixture was centrifuged at 400xg for 8 minutes with acceleration 9 and deceleration 9. All but 2 mL supernatant was removed and cleaning step was repeated. All but 1 ml supernatant was removed, remaining supernatant was resuspended and transferred to a clean 1.5 mL eppendorf tube. Sample was centrifuged at 1000xg for 8 minutes, supernatant was discarded, 500 µL PBS was added onto pellet and resuspended. Centrifugation was repeated and the supernatant was discarded. The pellet was immediately used for RNA isolation.

RNA isolation from PBMC samples was performed using RNeasy Mini Kit (Qiagen, USA). DNAase treatment mixture was prepared before, starting by mixing 1 µL DNase, 8 µL incubation buffer and 71 µL RNase water. The mixture was stored on ice until usage. For disrupting the pellet 600 µL RLT Buffer was added onto pellet and vortexed for 15 seconds. Equal volume of Buffer RLT, 600 µL 70% EtOH was added on lysate and mixed by pipetting. The mixture was transferred onto RNease spin column by dividing 2 times with 600 µL volumes and centrifuged at 8000xg for 15 seconds. The flow through was discarded at each step. Three hundred fifty µL RW1 Buffer was added into spin column and centrifuged at 8000xg for 15 seconds. The flow through was discarded, 80 µL DNAase treatment mixture was added into spin column and incubated for 15 minutes at room temperature. After incubation 350 µL RW1 Buffer was added into spin column and centrifuged at 8000xg for 15 seconds. The washing step was repeated with 500 µL RW1 Buffer. Second washing was performed with 500 µL RPE Buffer. The buffer was added and centrifuged at 8000xg for 15 seconds. Second washing was repeated with 2 minutes centrifugation. The spin column was transferred to a clean 1.5 mL Eppendorf and centrifuged at full speed for 1 minutes. Again the spin column was transferred to a clean 1.5 mL Eppendorf and 30-50 µL RNase free water depending on the cell pellet was added directly to the center of spin column membrane. It was incubated at room temperature for 1 minute and centrifuged at 8000xg for 1 minute. The RNA samples were stored at -80°C until further usage.

Quantification and quality control of DNA and RNA samples were completed immediately after isolations using NanoDrop 2000c Spectrophotometer (Thermo Fisher Scientific, USA). The ratio of 260/280 (1.8-2.0 nm) and 260/230 (2-2.2 nm) of DNA samples were considered for including to DNA sequencing analysis. Qubit® dsDNA HS Assay Kit (Thermo Fisher Scientific, USA) was used for quantification of double stranded DNA prior starting library preparation. For each sample and standard, the working solution was prepared by mixing 1 µL Qubit® dsDNA HS Reagent and 199 µL Qubit® dsDNA HS Buffer. Two µL of samples were mixed with 198 µL of working solution. Standard graph was recreated each week by mixing 10 µL of standard 1 and 2 with 190 µL working solution separately. Prior quantifying the standards and samples, each mixture was mixed by vortexing for 3 seconds and incubated at room temperature for 2 minutes.

3.3. Ion Personal Genome Machine™ Platform Targeted Next Generation Gene Sequencing

Sample number for each chip run was determined by using the following formula; *Total capacity of chip=Targeted coverage*sample number*size of primer panel*. Ion 318™ Chip Kit was used for experiments and total capacity of the chip was 2 Gb. Targeted coverage of our experiments was 500X and considering the final trimmings from bad quality reads our sample number per experiment was set as 32.

3.3.1. Primer design

Primer panel, containing *UNG* (chr12:109,534,879-109,548,796), *NEIL1* (chr15:75,637,831-75,647,592), *POLβ* (chr8:42,195,472-42,229,331) and *APOE* (chr19:45,408,011-45,412,655) genes including promoter, intron and exon regions was designed using Ion AmpliSeq™ Designer software (https://www.ampliseq.com). The primer panel was comprised of 226 amplicons which were divided into 2 equal pools. Primer covering regions, amplicon codes, pool placing and forward7reverse sequences were shown for each gene in Attachment 1. *POLβ* gene had the highest number of amplicons, and APOE had the lowest, proportional with the size of genes. The length of the amplicons was between 125-375 bp (Table 3.1). Total coverage of the primer panel was 94.5 ± 4.7 % and covered region's total size was 55.55 kb. *APOE* gene

primers had the highest coverage. Uncovered regions of primers were represented at Figure 3.1.

Table 3.1. Targeted next generation sequencing primers. The beginning and ending positions with targeted and covered base pairs are summarized. Also amplicon count and coverage percentage is included.

	POLβ **Chr8**	***UNG*** **Chr12**	***NEIL1*** **Chr15**	***APOE*** **Chr19**	***APOE*** **promoter**
Amplicon's beginning position	42,195,472	109,534,879	75,637,831	45,409,033	45,408,011
Amplicon's ending position	42,229,331	109,548,796	75,647,592	45,412,655	45,409,011
Targeted base pair	33859 bp	13919 bp	9761 bp	1220 bp	1000 bp
Covered base pair	29916 bp	13190 bp	8922 bp	1217 bp	980 bp
Amplicon count	124	54	36	8	4
Coverage Percentage	88.35 %	94.76 %	91.4 %	99.75 %	98%

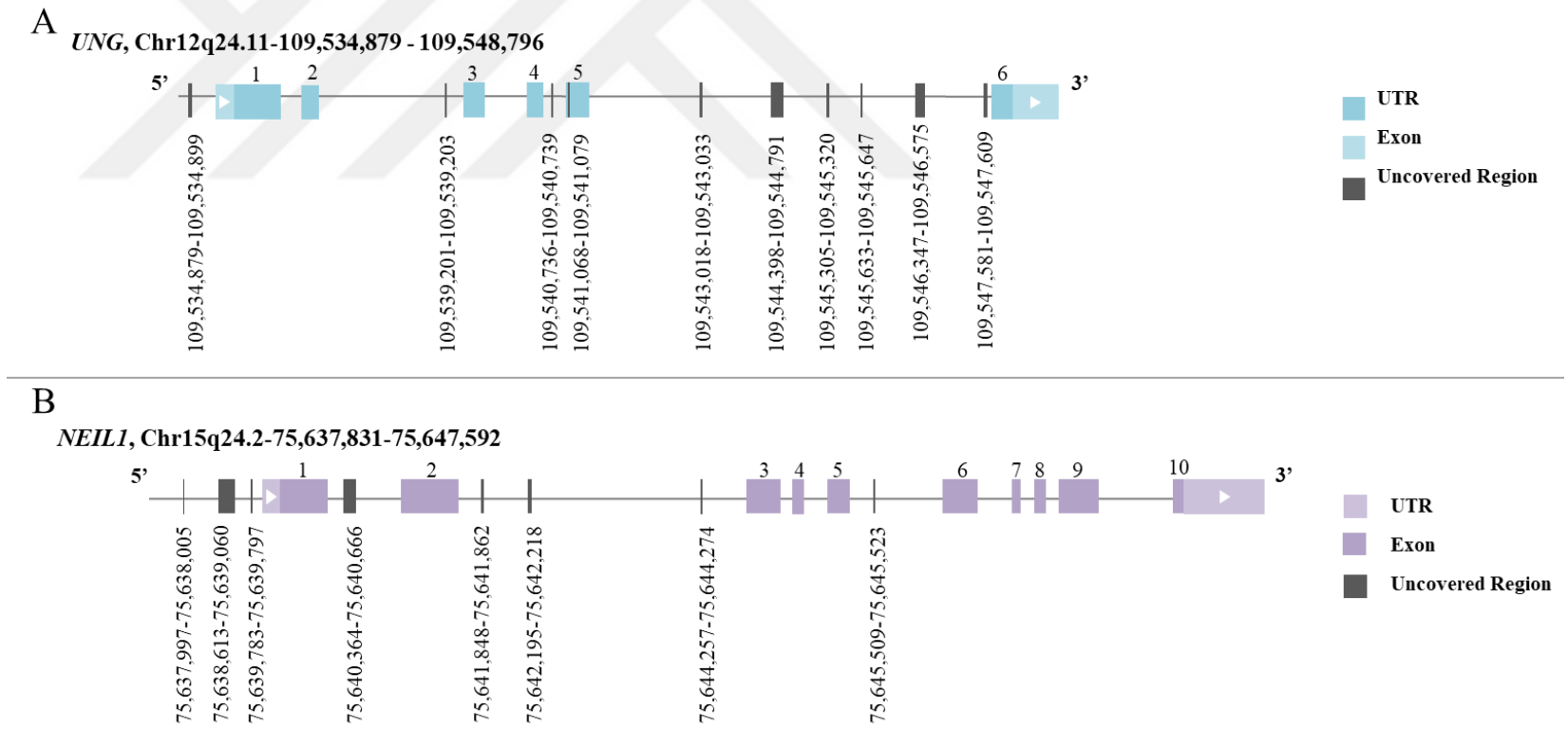

Figure 3.1. Gene structures of *UNG* and *NEIL1*, *POLβ* and *APOE.* A. *UNG* gene B. *NEIL1* gene C. *POLβ* gene D. *APOE* gene. Light colored boxes represent UTR; dark colored boxes represent exon regions. Uncovered regions of amplicons were represented with gray boxes

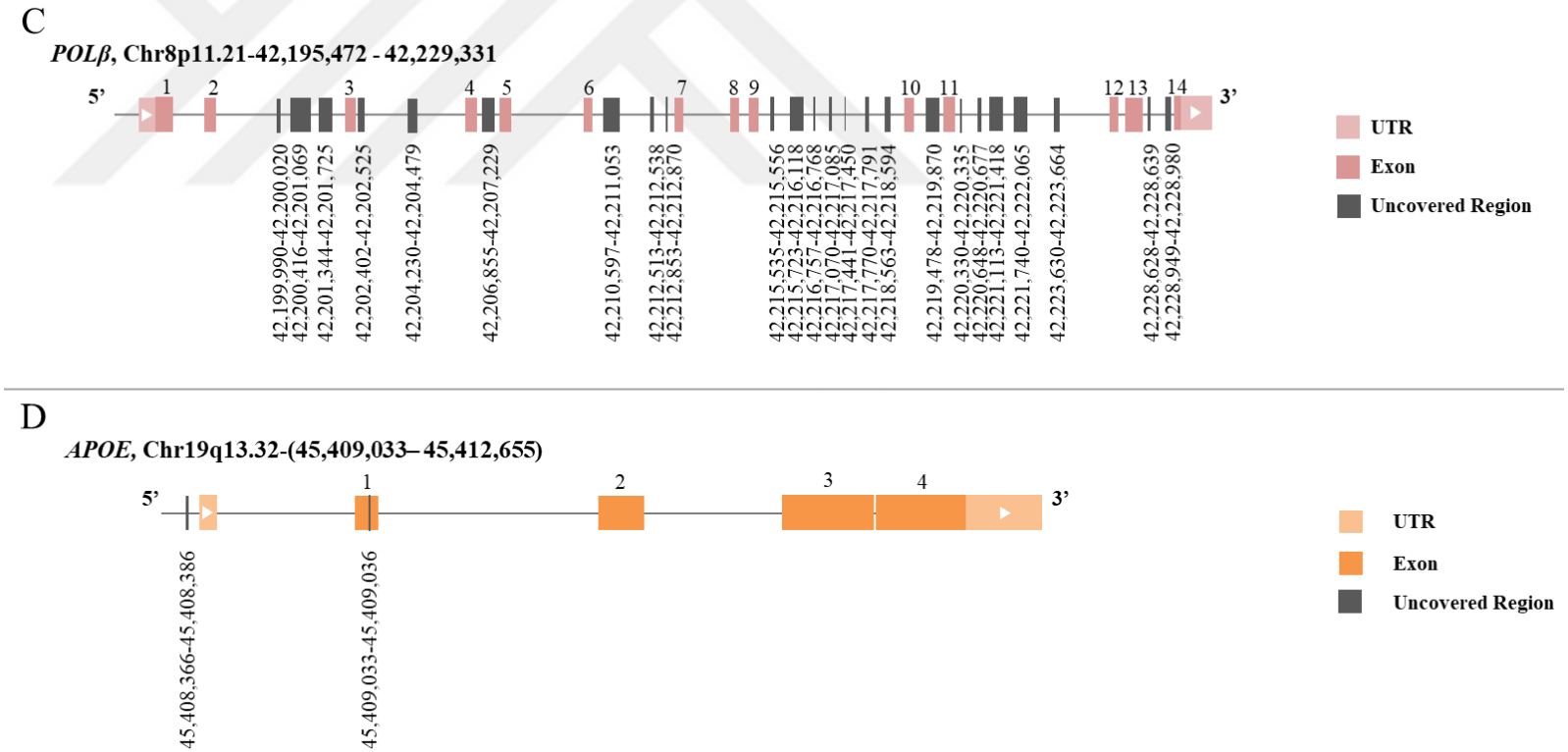

Figure 3.1. Gene structures of *UNG* and *NEIL1*, *POLβ* and *APOE* (continued). A. *UNG* gene B. *NEIL1* gene C. *POLβ* gene D. *APOE* gene. Light colored boxes represent UTR; dark colored boxes represent exon regions. Uncovered regions of amplicons were represented with gray boxes

3.3.2. Library preparation and equalization

DNA sample's concentrations were quantified and equalized to 3.4 ng/µL. Ion AmpliSeq™ Library Kits 2.0 (Thermo Fisher, USA) and Ion Xpress™ Barcode Adapters (Thermo Fisher, USA) were used for library amplification, preparation and sample barcoding.

DNA targets were amplified using multiplex PCR approach on Veriti™ 96-Well Thermal Cycler (Applied Biosystems, USA). Four µL of 5X Ion AmpliSeq™ HiFi Mix was mixed with 10 µL 2X Ion AmpliSeq™ Primer Pool for each sample. Primer pool-1 and pool-2 reaction mixtures were prepared separately. Six µL from DNA (~20 ng in total) aliquots were distributed to 2 different well in a 96-well plate which was divided as left and right side. Fourteen microliters of primer pool-1 mixture was added onto first well of the sample at the left side and primer pool-2 mixture was added onto other well of the sample at the right side as shown in Figure 3.2 and then pipetted up and down for 10 times.

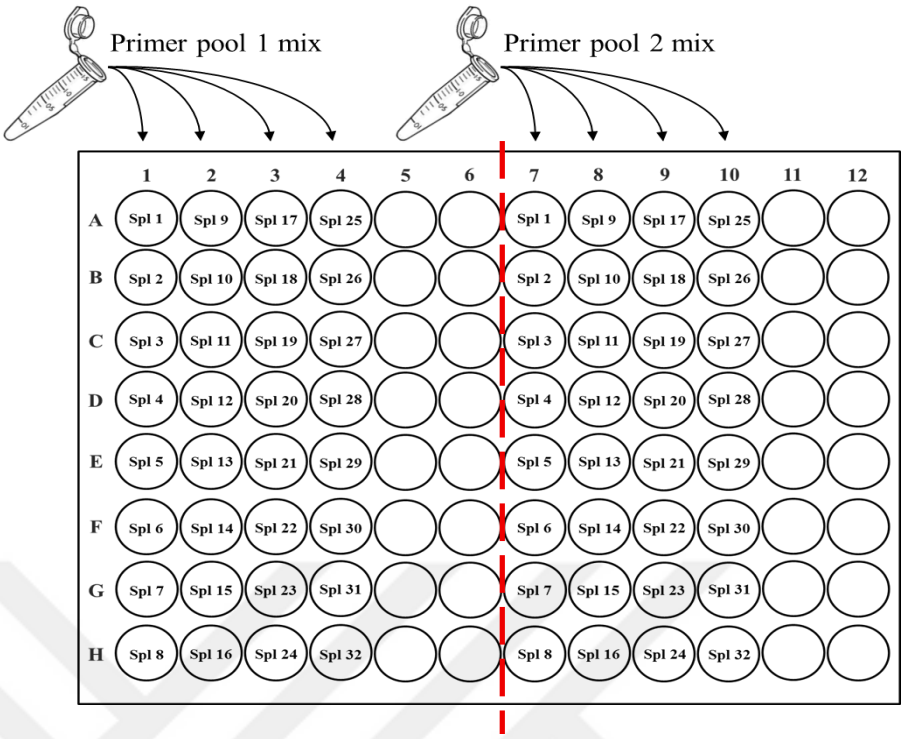

Figure 3.2. 96-well plate distribution. Plate was divided into 2 sides and samples were added into two different wells. Primer pool 1 mix was distributed to the left side and primer pool 2 mix was distributed to the right side of the 96-well plate.

Amplification was performed using following program on Veriti™ 96-Well Thermal Cycler (Applied Biosystems, USA); 2 minutes hold at 99°C following 21 cycle of 15 seconds denaturation at 99°C and 8 minutes annealing and extend at 60°C. Reactions were held at 4°C until next step or at -20°C for longer storage.

After amplification of samples, 2 µL FuPa reagent was added into samples for partial digestion and mixed by pipetting up and down for 10 times. Mixtures were incubated on Veriti™ 96-Well Thermal Cycler (Applied Biosystems, USA), first at 50°C for 10 minutes, 55°C for 10 minutes and finally at 60°C for 20 minutes.

Ion Xpress™ Barcode Adapters 1-16 and 17-32 Kits (Thermo Fisher, USA) were used for amplicon barcoding. Ion Xpress™ Barcodes were diluted by mixing following components; 2 µL Ion P1 Adapter, 2 µL on Xpress™ Barcode X and 4 µL nuclease

free water. Then 2 µL of diluted barcodes were mixed with 4 µL switch solution per sample. This mixture was distributed onto partially digested libraries and 2 µL of DNA Ligase was added and mixed by pipetting up and down for 10 times. Ligation reaction was incubated with following conditions on Veriti™ 96-Well Thermal Cycler (Applied Biosystems, USA); 22ºC for 30 minutes and 72ºC for 10 minutes.

Purification of barcoded libraries was performed using AMPure® XP reagent. 1.5X of sample volume, 45 µL in our case was distributed onto libraries and mixed by pipetting up and down for 10 times. The mixture was incubated at room temperature for 5 minutes and the 96-well was placed in DynaMag™-96 Side Magnet (Thermo fisher, USA) for 2 minutes incubation. After incubation the supernatants were discarded carefully without disturbing the pellets. For washing the pellets, freshly prepared 150 µL 70% ethanol was added onto pellets and the 96-well plate was moved side-to-side in the two positions of the magnet. The supernatants were discarded and washing step was repeated once more. After discarding the supernatant pellets were air-dried at room temperature for 5 minutes.

The 96-well plate was removed from magnet and the purified libraries were amplified adding, 50 µL of Platinum® PCR SuperMix High Fidelity and 2 µL of Library Amplification Primer Mix. Amplification was performed using following program on Veriti™ 96-Well Thermal Cycler (Applied Biosystems, USA); 2 minutes hold at 98ºC following 7 cycle of 15 seconds denaturation at 98ºC and 1 minute annealing and extend at 64ºC.

Amplified libraries were purified using AMPure® XP reagent. 0.5X of sample volume, 25 µL in our case was distributed onto libraries and mixed by pipetting up and down for 10 times. The mixture was incubated at room temperature for 5 minutes and the 96-well plate was placed into DynaMag™-96 Side Magnet for 5 minutes incubation. When the solution was completely clear, the supernatants were transferred to a clean 96-well plate. 1.2X of original sample volume, 60 µL in our case was

distributed onto libraries and mixed by pipetting up and down for 10 times. The mixture was incubated at room temperature for 5 minutes and the 96-well plate was placed into DynaMag™-96 Side Magnet for 3 minutes incubation. After incubation the supernatants were discarded. For washing the pellets freshly prepared 150 µL 70% ethanol was added onto pellets and the 96-well plate was moved side-to-side in the two positions of the magnet. The supernatants were discarded and washing step was repeated one more time. After discarding the supernatant pellets were air-dried at room temperature for 5 minutes. The 96-well plate was removed from the magnet and the pellets were resuspended with 50 µL of Low TE.

Amplified and purified library concentrations were quantified using Qubit® dsDNA HS Assay Kit as told in materials and methods 3.2. The concentrations of libraries were equalized to 22 ng/mL (~100 pM) using low TE. Diluted libraries were mixed in a 1.5 mL microcentrifuge tube with equal volume.

3.3.3. Template preparation

Clonal amplification of equalized libraries and enrichment of the template positive Ion PGM ™ Hi-Q™ Ion Sphere™ Particles were performed using Ion OneTouch™ 2 System and Ion PGM ™ Hi-Q™ OT2 Kit.

Ion OneTouch ™ 2 Instrument (OT2) was prepared prior to starting the emulsion based PCR step. Recovery Tubes containing 150 µL Ion OneTouch™ Breaking Solution were inserted to each slot in the centrifuge compartment. Recovery router was placed between the recovery tubes. Amplification plate was inserted into the heat block and the block was closed. The disposable tubing was first threaded through tubing catch and then secured in the pinch valve. The disposable injector was inserted and pushed through into injector hub port without touching anywhere until it reaches to the base of the router. New sipper tubes were placed to Ion OneTouch ™ Oil and Recovery Solution slots. One of the reagent tubes was filled half-full with oil and the other

reagent tube was filled quarter-full with recovery solution and they were placed to their corresponding slots. The waste container was emptied before starting the run.

Amplification solution was prepared, after the preparation of OT2 Instrument was completed. Ion PGM ™ Hi-Q™ Reagent Mix was allowed to come to the room temperature before using and it was vortexed for 30 seconds and centrifuged for 2 seconds. Ion PGM ™ Hi-Q™ Ion Sphere Particles (ISPs) was placed at room temperature. Mixed library was diluted to 8 pM by mixing 2 µL 100 pM library mixture and 23 µL nuclease free water. Diluted library was vortexed for 5 seconds and centrifuged for 2 seconds. Ion PGM ™ Hi-Q™ ISPs was vortexed for 1 minute at maximum speed and centrifuged for 2 seconds just before preparing amplification solution. Twenty five microliter of nuclease free water, 50 µL of on PGM ™ Hi-Q™ Enzyme Mix, 25 µL of diluted library mix and 100 µL of Ion PGM ™ Hi-Q™ ISPs were added onto pre-thawed Ion PGM ™ Hi-Q™ Reagent Mix. The amplification solution was mixed by pipetting up and down, vortexed at maximum speed for 5 seconds and centrifuged for 2 seconds.

Ion OneTouch™ Reaction Filter was placed into a tube rack and the amplification solution was added into the filter through sample port very slowly. One point seven milliliter Ion OneTouch™ Reaction Oil was added very slowly to filter through same sample port by dividing to two steps; 850 µL at each step. The filled reaction filter was installed to filter slot placing at the top stage of the OT2. Ion PGM™ Hi-Q™ OT2 Kit was selected from the drop down menu at the home screen for starting the run.

Template positive ISPs were recovered with final spin at the end of the run, if proceeding to enrichment step exceeded 15 minutes. After the centrifugation, recovery

tubes were removed and the recovery solution was removed all, but 100 µL from the tube without disturbing the ISP pellet. ISP pellets were resuspended in the remaining recovery solution and 500 µL of Ion OneTouch™ Wash Solution was added onto resuspension and pipetted up and down. Dispersed ISPs from each tube was combined in a new 1.5 mL eppendorf tube. ISPs were centrifuged at 15.500xg for 2.5 minutes and the wash solution was removed all but 100 µL.

Enrichment was started immediately after recovering step. Melt-off solution was prepared by mixing 280 µL Tween ® Solution and 40 µL 1 M NaOH. Dynabeads® MyOne™ Streptavidin C1 Beads were vortexed for 30 seconds for thoroughly re-suspending and 13 µL was transferred to a clean 1.5 mL eppendorf tube for washing. The tube was placed on DynaMag™-2 magnet and incubated for 2 minutes. After incubation the supernatant was removed and the pellet was resuspended with 130 µL MyOne™ Beads Wash Solution by vortexing for 30 seconds. The 8-well enrichment strip was loaded as shown in Figure 3.3 and inserted to Ion OneTouch™ ES Instrument, square shaped tab looking to left. A clean new tip was inserted in to tip arm and an opened 0.2 mL PCR tube, loaded with 10 µL of Neutralization Solution was inserted to tip loader hole. 8-well strip was pushed to the far-right position and the 35 minutes run was started. When the run was completed the tip and the 8-well strip was discarded. Enriched ISPs were gently mixed by inverting the tube for 5 times and stored at 4^{o}C until sequencing step was started.

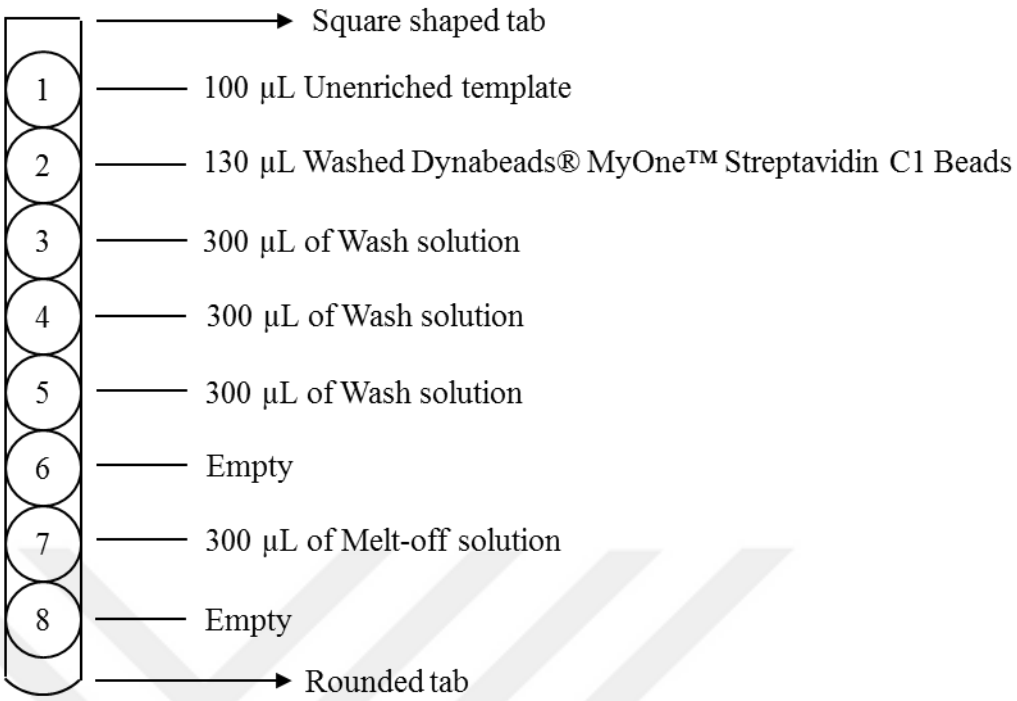

Figure 3.3. 8-well strip loading. 8-well strip was loaded with the unenriched template, beads, wash solution and melt-off solution in the following order for enrichment of the template.

Cleaning of the OT2 instrument was completed during the enrichment run. The left reagent tube was loaded half-full with Ion OneTouch ™ Oil if the volume was decreased below 20 mL. Reaction filter was discarded and Ion OneTouch ™ Oil was inserted to filter slot. Used injector of the amplification plate was inserted into a 50 mL falcon tube and cleaning step was started from the drop down menu.

3.3.4. Targeted gene sequencing

Next generation sequencing of the targeted libraries was performed using Ion PGM ™ Hi-Q™ Sequencing Kit and Ion Personal Genome Machine™ (PGM™) System (Figure 3.4). Plans of the each run were created from Torrent Suite Software on the Torrent Server which was connected to the PGM machine. Templates option from plan tab was selected and plan new run page was opened. From the tabs, application, kits, plugins, projects and plan options were defined according to starting material and kit type. Each sample's barcode number was also defined at the end of the run plan.

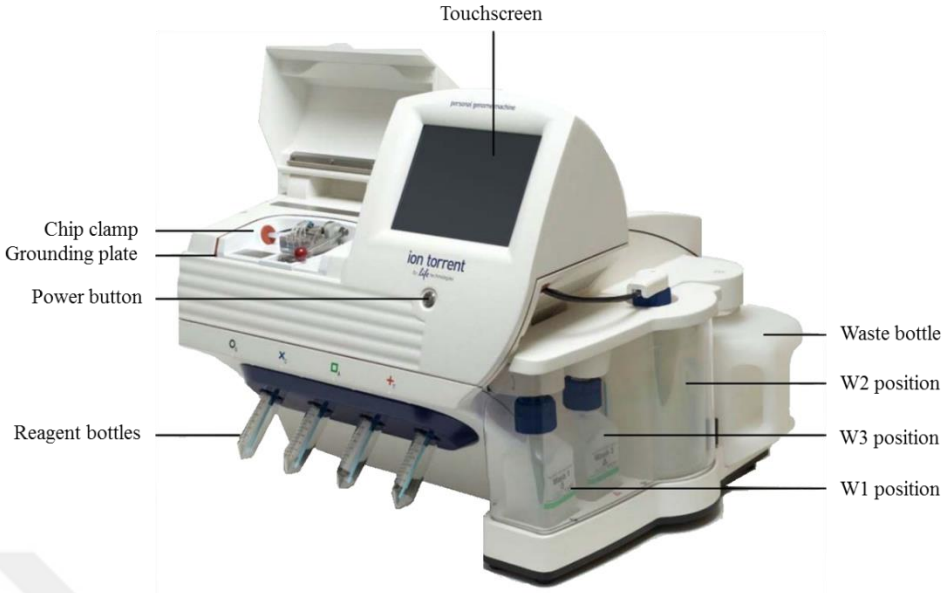

Figure 3.4. ION Torrent PGM System. PGM system and compartments were shown in the figure (Picture of the PGM system was taken from Ion PGM ™ Hi-Q™ Sequencing Kit protocol).

PGM system was an ion change based sequencing system which was using a semi-conductor chip for sequencing, placed under the chip clamp. The system was opened from the power button and controlled through the touchscreen. The ion sensitive layer beneath the chip compartment measured the pH changes during each flow of the dNTPs which were added into reagent bottles. Cleaning and the pH adjustment of the system was performed using specialized buffers which were added into the washing bottles and introduced through W1, W2 and W3 positions (Figure 3.4).

Wash 2 bottle from the Ion PGM™ Wash 2 Bottle Kit was conditioned before starting the experiments. The bottle was filled with 18 MΩ water to the mold line and the entire bottle of Wash 2 Bottle Conditioning Solution was added. It was inverted five times for mixing the solution and water, and incubated at room temperature overnight. After incubating the mixture was discarded.

At the beginning of each experiment PGM system was cleaned and prepared for the new run. Chloride cleaning was performed if the interval between two experiments

exceeded 1 week. Each cleaning bottle (2 250 mL and 1 2 L) was washed twice with 100 mL 18 MΩ water. Ion PGM™ Cleaning Tablet was completely dissolved in 1 L water and 1 mL freshly prepared NaOH was added. The solution was filtered using 0.22-µm filter and a 250 mL washing bottle was filled with filtered solution. Cleaning option was selected from the touchscreen and proceeded with chlorite cleaning option. Chlorite cleaning chip was secured. Old sippers on the PGM machine were used for cleaning steps and the sipper at the W1 position was washed with 18 MΩ water. The chlorite solution containing bottle was attached to W1 position and the cap was tightened. Empty 250 mL and 2 L bottles were attached to W3 and W2 positions in order. Collection trays were placed under the dNTP positions sipper and the cleaning was started by selecting next from the touchscreen.

Eighteen MΩ water cleaning was performed at the beginning of each run, independent from the experiment intervals. Each cleaning bottle was washed twice with 100 mL 18 MΩ water and 250 mL water cleaning water was filled with 18 MΩ water. 18-MOhm water cleaning option was selected from cleaning menu and the water cleaning chip was secured. W1 position sipper was washed with 18-MOhm water and water cleaning bottle was attached to the position. Empty 250 mL and 2 L bottles were attached to W3 and W2 positions in order. Collection trays were placed under the dNTP positions sipper and the cleaning was started by selecting next from the touchscreen. When the cleaning step was completed the sippers and bottles at the W1, W2 and W3 positions were removed and next option was selected for starting the initialization step.

Initialization was performed for pH adjustment of the PGM system. Wash 1, wash 2 and wash 3 bottles were washed three times with 18 MΩ water. Wash 2 bottle was filled with 18 MΩ water to the mold line and entire bottle of Ion PGM™ Hi-Q™ Sequencing W2 Solution was added. Seventy microliters of 100 mM NaOH was added into wash 2 bottle and the bottle was inverted 5 times for mixing thoroughly. Three hundred fifty microliters of 100 mM NaOH was added to the wash 1 bottle. Wash 3 bottle was filled with Ion PGM™ Sequencing W3 Solution to the 50 mL line. Initialize

was selected from the touchscreen and the PGM™ Hi-Q™ Sequencing W2 Solution bottle barcode was scanned. On the next step, the gas pressure was checked by PGM system. If the pressure of the system was enough, it would allow for proceeding to next step. Long gray sipper and prepared wash 2 bottle were attached to W2 position. Short gray sippers were attached to W1 and W3 positions immediately before attaching wash 1 and wash 3 bottles. The caps were tightened and the initialization was started by pressing next on the touchscreen. When the W2 solution pH reached 7.55 ± 0.1 sipper tubes at the dNTP positions and collection trays were removed and new blue sippers were attached to dNTP ports. Each reagent bottle was labeled as dGTP, dCTP, dATP, and dTTP and filled with 20 µL of corresponding dNTP stock solution. The reagent bottles were attached to corresponding dNTP ports and tightened. The initialization process was proceeded and the reagent bottles were filled with 40 mL W2 solution. At the end of the initialization step the pH of the system was checked and the chip loading was started if every reagent was at the target pH range.

Control Ion Sphere™ Particles was vortexed and pulse spinned. Five microliters of control ISPs were added onto enriched template and mixed by pipetting up and down. The mixture was centrifuged at 15.500xg for 2 minutes. All but 15 µL of the supernatant was discarded. Twelve microliters of thawed and vortexed sequencing primer was added onto the pellet and mixed by pipetting up and down. Primers were annealed by incubating at 95°C for 2 minutes and 37°C for 2 minutes. The reaction was stored at room temperature until the chip check was completed.

Ion 318™ Chip was used for sequencing which has up to 2 gb sequencing data memory (Figure 3.5A). The waste bottle was completely emptied prior starting the chip check. The next was selected from the touchscreen for proceeding to the following step. The cleaning chip was placed for cleaning the prompting lines and continued to next step. The OT2 option was selected as the template preparation instrument. The gloves were removed and hands were grounded from the grounding plate on the PGM system (Figure 3.4). The cleaning chip was replaced with the new chip. The new chip's barcode

was scanned and chip check was started. If the chip check was completed successfully the waste bottle was completely emptied and that option was selected from the touchscreen.

Sequencing polymerases were bind to the ISPs before loading to the chip. Three microliters from on PGM™ Hi-Q™ Sequencing Polymerase was added onto ISPs and incubated for 5 minutes at room temperature. The liquid was removed from the sequencing chip using a pipette with 45 degree tilt. For being ensure that all the liquid was removed, the chip was placed to the mini-fuge upside down with the chip tab pointing in and centrifuged for 5 seconds. The chip was placed into the bucket right-side up and placed on a firm, flat surface. Whole volume of the ISPs (~30 μL) was loaded from the loading port very slowly into the chip. The chip in the bucket was placed to mini-fuge with the chip tab pointing in and centrifuged for 30 seconds (Figure 3.5B). The centrifugation was repeated with the chip tab pointing out. The sample was pipetted out and in very slowly for one time by holding the chip in the bucket with a 45 degree. The liquid was discarded after the pipetting and placed into the chip bucket upside down. The chip was centrifuged for 5 seconds and any remaining liquid was discarded.

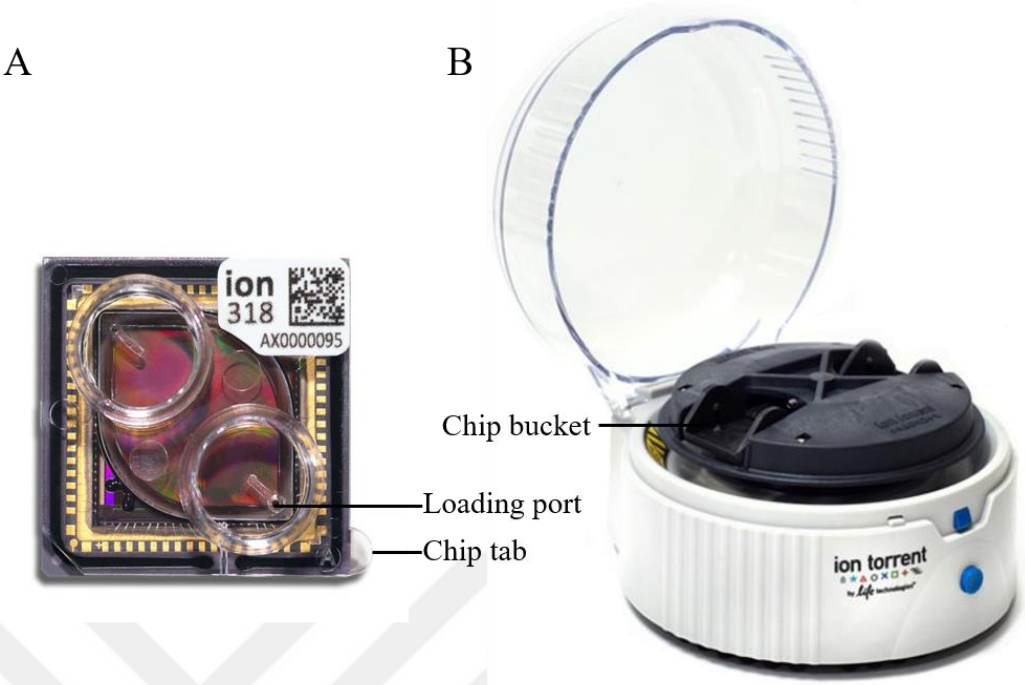

Figure 3.5. Ion 318™ Chip and minifuge A. Ion 318™ Chip loading port and chip tab was shown in the figure. B. Minifuge was used for centrifugation steps of the chip. (Picture of the PGM system was taken from Ion PGM ™ Hi-Q™ Sequencing Kit protocol).

The planned run name was selected from the touchscreen and next was selected. The settings were checked and proceeded to the next step. The loaded chip was placed and clamped. The run was started and the chip was visually inspected for any leak. After the chip calibration, sequencing was started by the system.

The 18-MOhm water cleaning step was performed after the run was completed before shutting down the system.

3.4. Bioinformatics and Statistical Analysis

Torrent Suite Software v5.0.4 plugins (Thermo Fisher Scientific, USA) was used for bioinformatics analysis of the first generated raw data. The FASTQ files as the result of Ion-PGM system were trimmed with the qualified standards of the system to generate unaligned BAM (UBAM) files. BAM files were created by aligning the

UBAM files to the GRCh37-hg19 human reference genome. Variant calling step for VCF files was performed using Variant caller plugin.

Ion Reporter Software (Thermo Fisher Scientific, USA) was used for the location, zygosity, position, type analyses and accession number of the variations of VCF files. CLC Genomics Workbench (9.0.1., Qiagen, USA) was used for the comparative bioinformatics analyses of case-control groups using the VCF files of samples. CLC Genomics Workbench (9.0.1., Qiagen, USA) was also used for each dataset's quality assessment of the sequencing results. Bonferroni corrected Fisher's exact test was used for selecting statistically significantly important variations according to the p-value. Pearson $\chi 2$ test and odds ratio (OR) with 95% confidence interval (CI) were used for supporting the Fisher's exact test results. Also Hardy-Weinberg equilibrium (HWE) was applied to each variation among controls. Integrative Genomics Viewer (IGV) tool was used for the integrity control of the statistically significantly important variation's amplicons (104).

Dominant, recessive and additive genetic models were tested for the statistically significantly important variations. G*Power software version 3.1.9.4 (Institute for experimental psychology in Dusseldorf, Germany) was used for calculating the statistical power of statistically significantly important variations (105). Haploview 4.2 (Broad Institute of MIT and Harvard, Cambridge, MA, USA) was used for performing Linkage Disequilibrium (LD) and haplotype analysis of statistically significantly important variations.

3.5. Sanger Sequencing

Variations that has been found statistically significantly different, was conformed using Sanger sequencing.

3.5.1. Primer design and initial PCR

Primer designs of the Sanger sequencing for *POLβ, UNG, NEIL1* and *APOE* genes were performed using Primer Blast, NCBI primer finding tool. Homo sapiens

genome sequence, hg19 was set as the reference genome. Primers matching with only intended regions were selected. The length of the primers was maximum 600 bp. Primer codes, covering variations, forward and reverse primer sequences and annealing temperatures were listed in Table 3.2.

Table 3.2. Sanger sequencing primers of *POLβ, UNG, NEIL1* and *APOE*. The locations included in primers, forward/reverse primers and their annealing temperatures were summarized.

Primer ID	Location	Forward primer/ Reverse primer	Annealing temperature
P1	42,197,233	CACATCCACCTGCTTCTACTCCTT/ CAAGAAGATTCACCATAAAGTG	55°C
P2	42,201,087	CAGCCTGGGTGACTCCAT/ CTGGCCTACGTCTAACTTT	56°C
P3	42,201,998 42,201,999	ATCCTAGCTGGTTCACACCTGT/ CTGGCCTAATTTTCTACCTCCA	58,5°C
P4	42,204,931	AGTTGCTATCTAGCCCATTTC/ AAGCAATTCAGACCCAAGCTAC	55°C
P5	42,216,968 42,216,970 42,217,057	CTCTAAGACATCTCACTTCCA/ CACTTGAGCCAAGAGTTTGA	60°C
P6	42,219,877	CAGCCTGGCAACAGAACGAG/ GAGACTGTGCTATTCACAGGTACA	62°C
P7	42,220,369 42,220,387 42,220,396	GCCATTCTGAGTGTGTGACTTC/ AAGAGCTGGGACTACACCACCA	62°C
P8	42,220,710 42,221,065 42,221,070	TAGTTAATTAGGCCTGGGCATGG/ TAGGGAAAAACTTACTGTGTGTCAA	60°C
P9	42,221,455 42,221,461 42,221,463	TGTGATAGAACACTCAGGGGA/ GAATCTTAATTCTAGGCTGGGCA	58°C
P10	42,227,356^ 42,227,357	TTACAAGGTTGGGGTGATT/ AACTACACTAACTGGCATAG	55°C
P11	42,203,655	TTGGATTAGTGATACTCAAC/ ACCTGGACTACTCAAAAGG	54°C
P12	42,227,787	AACTATGCCAGTTAGTGTAG/ CTAGTCTACCAGGGATGA	55°C
P13	42,204,068 42,224,089^ 42,224,090	CTGAGATGCCTTTTGAGTAGT/ CAGTTATGGTCCAGTCATGG	62°C
P14	42,204,551^ 42,204,552	TCTTGGTGAACTGACCCATTGA/ GAAATGGGCTAGATAGCAACTC	58.5°C
P15	42,223,394	TGACCAGGAAGGTAGTTCCA/ CAGGTTACATCCAGCCCACTA	58.5°C
P16	42,207,347	ATGTTTACACTCTCCCAAGCAT/ AGTCTGCTAAAAGACATGAAGA	58.5°C
U1	109,535,809 109,535,857	GCGCCTCTGACTCGGTAAAC/ CGTGGTCTTTTCGCGGCA	58.5°C
U2	109,540,276	ATCCCTGGCCTCTACCAC/ GGATTTAACAGCACTCTGGA	58.5°C
U3	109,537,423 109,537,562	TCCCGCGGTGTCAGATGTT/ ACCAGGTTTCATTGGTCTTGAACTC	58.5°C

U4	109,545,819 109,545,820 109,545,951	GGCACAAAATGGAAAGTAAATGGC/ TCCCAGCCAAGAGACCCTG	58.5°C
U5	109,546,114	GTACAACTGGTCCCTGGCTTA/ CCACCTCCTAGGCTCAAGTTAC	58.5°C
U6	109,546,886	CTCTTGGTCCCATGCAGTT/ TGCAAAGAGCACCACCCTC	58.5°C
U7	109,548,571^ 109,548,572	TTGTTATTCCTTGGGTGTTGG/ GGTGCAGACCCAATAAAG	55°C
U8	109,541,004 109,541,007 109,541,008 109,541,009 109,541,043 109,541,051.. 109,541,052	GTGGGCCAAGCAAGGTAA/ AGGCAAATCAGCCCAATACC	60°C
U9	109,541,435^ 109,541,436 109,541,462^ 109,541,463	CTGGGTATTGGGCTGATTTGC/ CCAGTTGCTTTGCTTTCCTC	60°C
U10	109,541,171	GCAGCAGGGAAGCACAGC/ CAAACACAGCAGGGACTCCTA	62°C
N1	75,643,714	ATGGCAAACCCACATATGGTC/ GGCAACAAGGAAGACCCCAT	62°C
N2	75,645,207 75,645,209 75,645,238 75,645,383	GCACTTGTCCCTCTCTGGA/ CTCATGGGCAGCTTGGAGGA	62°C
N3	75,645,565 75,645,566.. 75,645,567	ACAGTGTTTCCTCCAAGCTG/ TCCTGGAGTGGGATTAAGAA	58.5°C
N4	75,642,571.. 75,642,572	CCTGCCTGCAGTAGGGTGTA/ TCCATTTCTCATTCCCCTATCC	58.5°C
N5	75,645,965	CCTGGACTAGCCTCAAACTG/ CAGAGGGCAGGCTCAGGTT	62°C
N6	75,644,183	GCCACCACACCAGCCTCAAG/ TCTGTTAAATGGGAACATCAAGGGA	62°C
A1	45,410,002	ACAGGCAGGAAGATGAAGT/ CCAGAGAGCGTCAAATCGCT	58.5°C
A2	45,411,941	AGCTCCTTCTTCGTCTCTGC/ CACTGCCAGGCGCTTCTG	62°C

Initial PCR of Sanger sequencing was performed using two different taq polymerases. First one was; Taq DNA Polymerase, recombinant (Invitrogen, USA) and dNTP mix 10mM ea (Thermo Fisher, USA) and the second one was; Platinum™ SuperFi™ DNA Polymerase (Thermo Fisher, USA). Both PCR reactions was performed using Veriti™ 96-Well Thermal Cycler (Applied Biosystems, USA).

PCR master mixtures of 1st polymerase were prepared using 1X PCR Buffer, $MgCl_2$, 0.2 µM dNTP mixture 0.25 mM forward and reverse primers and 2U Taq DNA

Polymerase in final concentrations and mixed with 50 ng DNA sample. PCR conditions were as follows; 2 minutes hold at 94°C following 35 cycle of 45 seconds denaturation at 94°C and 30 seconds annealing at specific temperature for each primer and at specific durations for each primer extension at 72°C and finally a final extension for 10 minutes at 72°C. Reactions were held at 4°C until next step or at -20°C for longer storage. Primers that were amplified using 1st polymerase, their specific $MgCl_2$ concentrations and extension durations were given in Table 3.3.

Table 3.3. First polymerase primers. Extension durations and MgCl2 concentrations of the 1st polymerase reactions**.**

Primer ID	Extension duration (seconds)	$MgCl_2$ concentration (mM)
P4	30	1
P5	60	1.5
P6	30	1
P9	30	1
P10	30	1
P11	60	0.5
P15	30	1
P16	30	1
U3	60	1.5
U4	30	1
U6	30	1
U10	30	0.75
N1	60	0.5
N3	30	1
N4	30	1
N5	30	1
N6	30	0.75
A1	30	0.75
A2*	60	0.75

*DMSO was added to te reaction mixture to enhance specific binding

PCR master mixtures of 2nd polymerase were prepared using 1X SuperFi Buffer, 0.2 mM dNTP mix, 0.5 μM forward and reverese primer optional 1X SuperFi GC Enhancer and 0.02 U Platinum SuperFi DNA Polymerase. PCR conditions were as follows; 2 minutes hold at 95°C following 35 cycle of 10 seconds denaturation at 95°C and 10 seconds annealing at specific temperature for each primer and 20 seconds extension at 72°C and finally a final extension for 5 minutes at 72°C. Reactions were held at 4°C until next step or at -20°C for longer storage. Primers that were amplified using 2nd polymerase were given in Table 3.4.

Table 3.4. Second polymerase primers. Primers that were amplified with second polymerase and their GC enhancer usage situation.

Primer ID	GC Enhancer usage
P1	-
P2	-
P3	-
P7	-
P8	-
P12	-
P13	-
P14	-
U1	-
U2	+
U5	+
U7	-
U8	-
U9	-
N2	+

PCR products were run on agarose gel. Tris-Acetate-EDTA (TAE) buffer was prepared for electrophoresis. For 50X TAE Buffer; 242 grams of Tris base, mw: 121.14 g/mol (Sigma Aldrich, USA) was dissolved in 0.6 L dH_2O by stirring on a stirrer and after it completely dissolved, 18.61 g Disodium EDTA, mw: 336.21 g/mol (Sigma Aldrich, USA) was added into H_2O. The mixture was stirred on a stirrer overnight. After the solution was completely homogeneous 57.1 mL 100% acetic acid (Sigma Aldrich, USA) was added onto the solution. The volume of the solution was completed to 1L. For 1.5% agarose gel preparation 1X TAE buffer was prepared from 50X by mixing 20 mL 50X buffer with 980 ml dH_2O. 2.25 g Agarose, for molecular biology, low eeo (Sigma Aldrich, USA) was weighed and mixed with 1X TAE buffer and heated in a microwave for completely melting the agarose. After the agarose was melted and the temperature was decreased to 60°C 0.5 μg/mL EtBr (Sigma Aldrich, USA) was added into the mixture and mixed until it homogeneously dissolved. The mixture was poured into 150 mL gel tank of Midi plus-2 Horizontal Electrophoresis System (Major Science, USA) and waited until the gel completely solidified. The gel was run in 1X TAE for 1.5 hours at 100 V and visualized using ChemiDoc Imaging System (Bio-Rad, USA).

PCR products were purified using ExoSAP-IT™ PCR Product Cleanup Reagent (Applied Biosystems, USA). 5 µL from PCR products were mixed with 2 µL ExoSAP reagent. The mixture was incubated first at 37°C for 15 minutes and then at 80°C for another 15 minutes on Veriti™ 96-Well Thermal Cycler (Applied Biosystems, USA).

3.5.2. Final Sanger sequencing PCR and cleanup

Cycle sequencing was performed using BigDye™ Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, USA). The reaction mixture was prepared using 2 µl BigDye™ Terminator 3.1 Ready Reaction Mix 2 µL BigDye™ Terminator v1.1 & v3.1 5X Sequencing Buffer, 2 µL 2 µM forward or reverse primer, 2,5 µL dH2O and 1,5 µL purified PCR product. PCR conditions were as follows; 1 minutes hold at 96°C following 25 cycle of 10 seconds denaturation at 96°C, 5 seconds annealing at 50°C and 4 minutes extension at 60°C. Reactions were held at 4°C until next step.

Final Sanger sequencing PCR product cleanup was performed using BigDye XTerminator™ Purification Kit (Applied Biosystems, USA). Before using BigDye XTerminator™ bead solution it was vortexed for 10 seconds. SAM/BigDye XTerminator™ bead working solution was prepared by mixing 45 µL SAM solution and 10 µL BigDye XTerminator™ bead solution per sample. In a clean 1.5 mL eppendorf tube 10 µL PCR product was mixed with 55 µL working solution and the mixture was vortexed for 30 minutes at 2000 rpm on MS 3 Digital Shaker (IKA, China). After vortexing the mixture was centrifuged for 5 minutes at 1000xg. Supernatant was transferred to a clean 1.5 mL eppendorf tube.

3.5.3. Cycle sequencing

Sanger sequencing of the samples was performed using ABI PRISM® 310 Genetic Analyzer (Thermo Fisher, USA).

The instrument was prepared prior to sequencing. First the instrument was opened and then the computer was turned on. The run conditions and instrument was controlled from 310 Data Collection Software (Thermo Fisher, USA) on computer. The glass

syringe was released from the program and was filled with 0.6 mL POP-6™ Polymer for 3500/3500xL Genetic Analyzers (Applied Biosystems). The syringe was placed to its holder and sealed with computer command. The polymer was passed through the channels for removing the air bubbles prior sequencing.

310 and 31xx Running Buffer, 10X (Applied Biosystems) was diluted to 1X by mixing 1.5 ml 10X buffer with 13.5 ml dH_2O for preparing working buffer. The anode buffer vial was filled with 1X buffer up to red line. The number 1 glass vial placed in automatic sampler tray was filled with 1X working buffer and number 2 glass vial was filled with dH_2O up to red line. An empty 1.5 mL Eppendorf tube was placed to the 3rd tube section.

A new run was created on 310 Data Collection Software. The type of the polymer, sample preparation kit and the locations of the samples were selected for the run.

For sequencing 15 µL from purified PCR samples was transferred to clean 0.2 mL PCR tubes. The samples were placed to the pre-defined locations and the run was started from the software. ABI PRISM® 310 Genetic Analyzer software and FinchTV v1.5.0 (Geospiza, Seattle, WA).

3.6. Reverse Transcription and Quantitative Real Time PCR

Reverse transcription reactions of the RNA samples were performed using iScript™cDNA Synthesis Kit (Bio-Rad, USA). For cDNA conversion 500 ng from RNA samples was used. For reaction mixture 4 µL 5X iScript Reaction mix and 1 µL iScript Reverse Transcriptase was mixed for each sample. Nuclease free water volume was calculated according to the RNA sample volume and total volume was completed to 20 µL. The reaction mixture was mixed with RNA samples and incubated at 25°C for 5 minutes, at 46°C for 20 minutes and at 95°C for 1 minute. cDNA samples were stored at -20°C until gene expression analyses.

For UNG gene expression analysis quantitative real time PCR (qRT-PCR) was performed, using PowerUp™ SYBR™ Green Master Mix (Applied Biosystems, USA) LightCycler® Nano (Roche, Switzerland) instrument. Following primers were used for *UNG* gene: forward primer: 5'-ACTGTTTATCCACCCCCACAC-3' and reverse primer: 5'-ATGTTCTCCAAACTGGGCGGA-3'. *β-Actin* was used as a housekeeping gene:; forward primer: 5'-CACCAACTGGGACGACAT-3' and reverse primer: 5'-ACAGCCTGGATAGGCAACG-3'.For reaction mixture 10 µl 2X PowerUp™ SYBR™ Green Master Mix, 0.6 µL 300 nM forward and reverse primer (from 10 µM stock) and 6.8 µL dH_2O was mixed. The reaction mixture was mixed with 2 µL cDNA sample. The reaction conditions were given in Table 3.5. UNG mRNA relative amount was calculated by first normalizing the UNG Ct value against β-actin reference gene Ct-value. Normalized expression levels of cases and control were compared using standard ΔCT and $2^{-\Delta\Delta CT}$ method. The significance was determined by using Student's t-test (Mann-Whitney test) in GraphPad Prism 5.03 and 0.05 was set as threshold p-value.

Table 3.5. qRT-PCR reaction conditions of gene expression analysis**.**

Step	**Temperature**		**Duration**
UDG activation	50°C		2 minutes
Dual-Lock™ DNA polymerase	95°C		2 minutes
Denature		95°C	15 seconds
Anneal	40 cycles	58°C	15 seconds
Extend		72°C	1 minute
Melting curve stage			
Ramp rate	**Temperature**		**Duration**
1.6°C/second	95°C		15 seconds
1.6°C/second	60°C		1 minute
0.15°C/second	95°C		15 seconds

4. RESULTS

4.1. Study Population

The targeted gene sequencing of 198 LOAD and 48 sporadic PD patients and 98 control individuals were completed. The mean age of the LOAD, control and sporadic PD groups are as follows; 79.85±7.83, 74.04±7.62 and 71.23±7.24. In LOAD and control group the females were more than the males but not in sporadic PD. In LOAD-blood group patients with moderate and severe MMSE scoring were almost equal while most of them had severe CDR scoring. In sporadic PD group most of the patients were in stage 2 Hoehn and Yahr scale. The age, female/male ratio, MMSE and CDR scores and Hoehn and Yahr Scale of the patients and controls were summarized in Table 4.1.

Table 4.1. Study population demographics of blood samples. The means of ages, female/male ratios and scorings of LOAD and sporadic PD are summarized**.**

Characteristics	LOAD, n=198	Control, n=98	Sporadic PD, n=48
Age, mean ± SD[a]	79.85±7.83 (range: 65-97)	74.04±7.62 (range: 65-96)	71.23±7.24 (range:53-80)
Female/Male	118/80	57/41	17/31
MMSE[a] score			
>20, mild (n)	21.38±1.77 (51)		
10-19, moderate (n)	15.13±2.56 (74)	Normal	N/A
<10, severe (n)	6.71±2.41 (73)		
CDR[a] score	**n**	**n**	**n**
0, normal	0	98	
1, mild	49	0	N/A
2, moderate	64	0	
3, severe	85	0	
Hoehn and Yahr Scale	**n**	**n**	**n**
Stage 0			0
Stage 1			12
Stage 1.5			10
Stage 2	N/A	N/A	16
Stage 2.5			2
Stage 3			7
Stage 4			1
Stage 5			0

[a]MMSE, mini-mental state examination; CDR, clinical dementia rating scale; SD, standard deviation; N/A, not applicable for related group

In addition to the blood samples, the targeted gene sequencing of postmortem brain tissues of LOAD (10 TC and 11 CE), cognitively normal controls (9 TC and 10 CE) and hpC subjects (8 TC and 11 CE) were completed. The CE and TC samples belongs to the same individuals. The mean age of the LOAD, hpC and CT are as follows;

82.72±6.32, 83.27±5.71 and 76.8±7.52. In LOAD and hpC group the female to male ratio is high but in CT group they are equal. Braak stage V-VI and CDR score 3 patients are the dominant group in LOAD. In hpC there are more Braak stage III-IV individuals present and as expected all CDR scores of the both hpC and CT group are 0. The Braak stage of CT group is centered in stage I-II. The age, female/male ratio, Braak stage and CDR scores of the patients and controls were summarized in Table 4.2.

Table 4.2. Study population demographics of post-mortem brain tissue samples. The means of ages, female/male ratios and scorings of LOAD are summarized.

Characteristics	**LOAD, n=11**	**hpC, n=11**	**CT, n=10**
Age, mean ± SD[a]	82.72±6.32 (range: 72-94)	83.27±5.71 (range: 73-95)	76.8±7.52 (range:64-89)
Female/Male	8/3	8/3	5/5
Braak stage	**n**	**n**	**n**
0	0	0	2
I-II	0	0	8
III-IV	0	8	0
V-VI	11	3	0
CDR[a] score	**n**	**n**	**n**
0, normal	0	10	10
0.5, questionable	0	0	0
1, mild	0	0	0
2, moderate	2	0	0
3, severe	9	0	0

CDR, clinical dementia rating scale; SD, standard deviation; N/A, not applicable for related group

4.2. NGS Study Set Up Evaluation

Using the CLC Genomics Workbench software the basic quality statistics of the study were determined. More than 95% of the total 39,728,454 reads had average PHRED quality score 20 (Figure 4.1). The readings showed no ambiguous bases. The read length distribution fitting to the 125-375 bp interval was 64.5%.

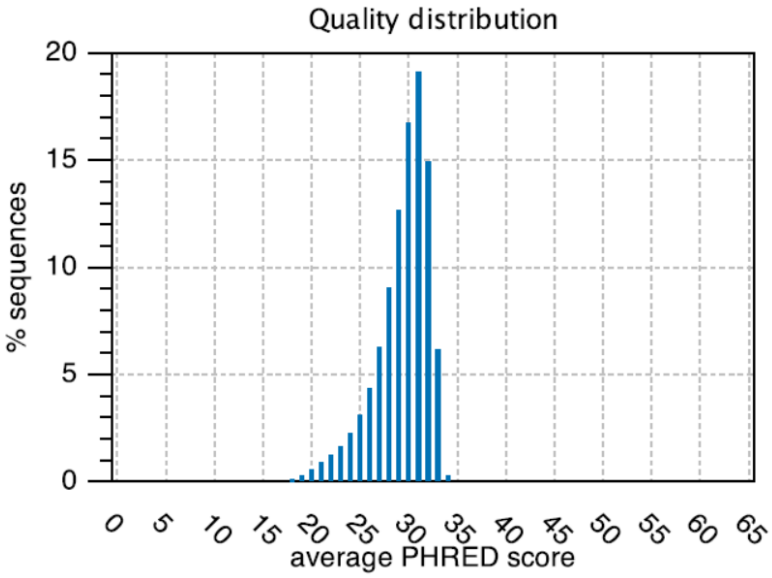

Figure 4.1. Average PHRED quality score distribution. Sequence percentage was compared against average PHRED score.

The total variations found in study groups at the end of the study were as follows; 907 in LOAD, 360 in sporadic PD and 544 in control blood samples; 403 in CE and 307 in TC of LOAD; 332 in CE and 282 in TC of hpC; 331 in CE and 328 in TC of cognitively controls. The distribution of these variants among studied; *UNG*, *NEIL1*, *POLβ* and *APOE* genes are shown in Figure 4.2A with their percentages. The highest percentage in the studied genes is belongs to *POLβ* in all studied groups followed by; *UNG*, *NEIL1* and *APOE* proportional to the gene sizes. The common variants of LOAD blood, CE and TC samples and also hpC CE and TC samples were summarized in Figure 4.2B for better understanding the distribution of variants according to LOAD pathology. All LOAD samples including asymptomatic hpC samples carry 50 common variants. Most common variations (19) between LOAD and post-mortem brain tissues are between LOAD-blood and LOAD CE samples. Between LOAD-blood samples and PD samples, there are 341 common variants. The highest percentage of these common variants are belong to *UNG* (57.8%) followed by *POLβ* (26.1%), *NEIL1* (10.1%) and *APOE* (5.3%).

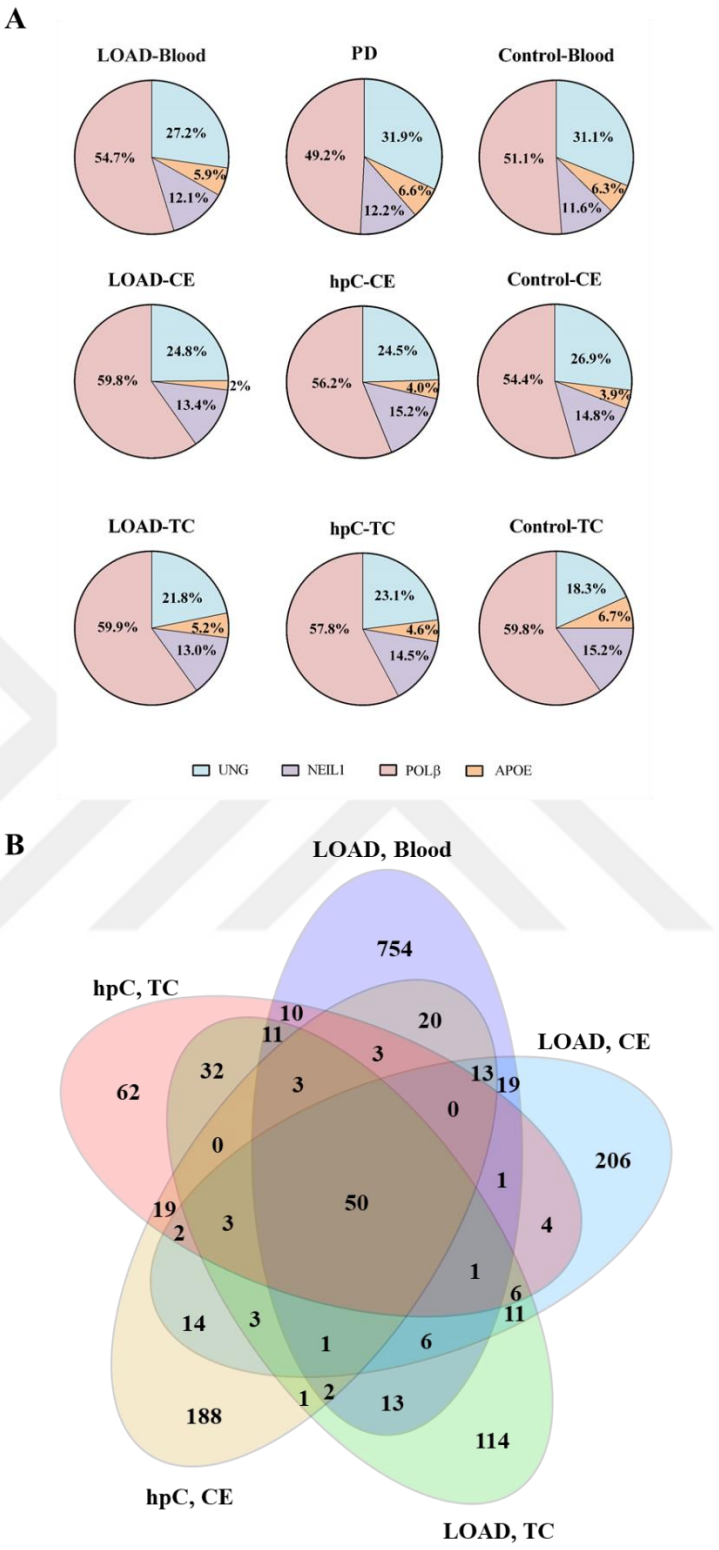

Figure 4.2. Genetic variant distributions A) Percent distribution of *UNG*, *NEIL1*, *POLβ* and *APOE* gene variants are represented as pie charts B) shared or specific LOAD and control variants are represented in venn diagram.

4.3. Comparative Analysis of Case and Control Groups

CLC Genomics Workbench Fisher's Exact test was used for performing comparative analysis of case and control groups. The results were ranked according to their p-value and for selecting the candidate variations for LOAD pathogenesis and following criteria were applied for selection; p-value < 0.05 with the exception of location of the variation and minor allele frequency of the variants in control group is smaller than 0.10. Thirty one candidate variations were selected for LOAD pathogenesis, 72% was SNVs and remaining was INDELs. The gene distribution of the selected BER genes are; *UNG*, 54.8%; *NEIL1* 12.9% and *POLβ*, 32.3%.

4.4. Sanger Sequencing Results

Candidate gene's confirmations were completed using both IGV tool and Sanger sequencing. The variants which are in close-range were also included to primer list. In total 32 primers were designed for 56 BER genes locations (20 *UNG*, 10 *NEIL1* and 25 *POLβ)*, 32 candidate and 24 close-range variations. *APOE* rs769449 variant showed one of the strongest interactions so it was also confirmed along with the BER variants.

4.4.1. Positive results

Thirty percent of the variants were confirmed as positive (17 in total; 8 *UNG*, 5 *NEIL1* and 3 *POLβ)*. *UNG* gene variants showed highest positive rate among other genes. Forty percent of the *UNG* gene variants (rs1018782, rs1018783, rs80001089, rs2430678, rs2515913, rs2268406, rs2268406) were confirmed as positive, including 1 insertion (rs1610925). Half of the selected *NEIL1* variants (rs11634109, rs767369942, rs5745916) were also confirmed, including 1 insertion (rs10653888) and 1 deletion (rs5745925). Only 12% of the *POLβ* gene variants were confirmed as positive, which are all SNPs (rs11993638, rs571459229, rs3136744). The included *APOE* variant rs769449 was also confirmed as positive. The IGV presentations with reading counts and Sanger sequencing chromatograms of the positive results are shown in Figure 4.3.

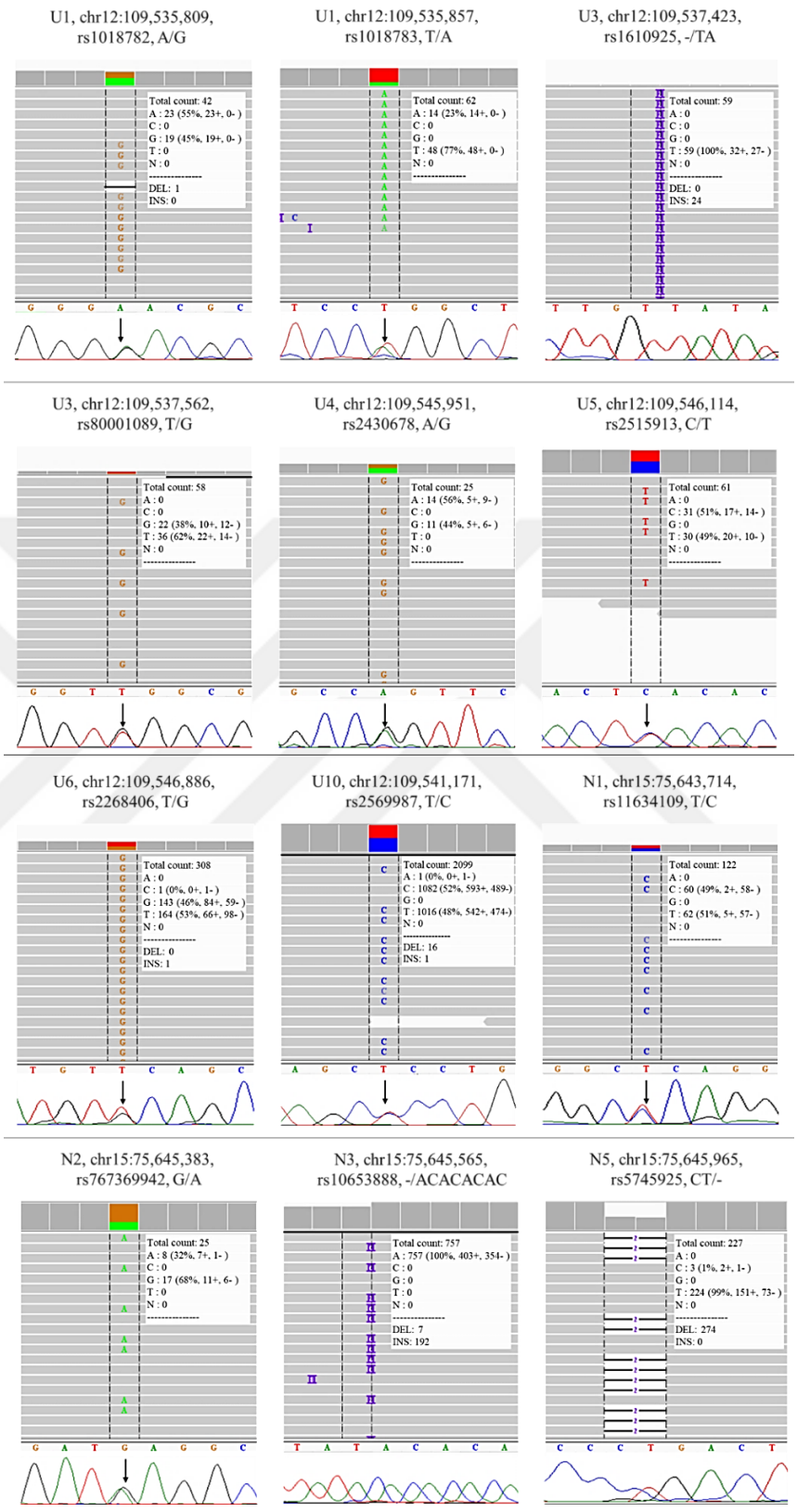

Figure 4.3. Positive Sanger sequencing results. IGV presentations are aligned with Sanger sequencing chromatograms. The reading counts of the variants are also included U, *UNG*; N, *NEIL1*; P, *POLβ*; A, *APOE*.

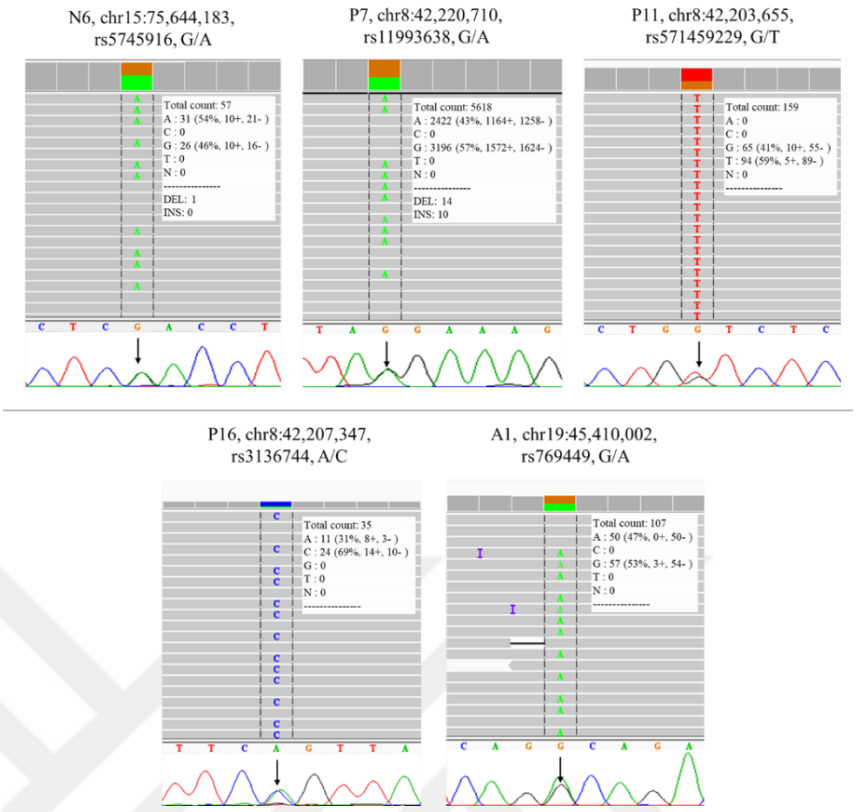

Figure 4.3. Positive Sanger sequencing results (continued). IGV presentations are aligned with Sanger sequencing chromatograms. The reading counts of the variants are also included U, *UNG*; N, *NEIL1*; P, *POLβ*; A, *APOE*.

4.4.2. False positive results

Seventy percent of the selected variants were found to be false positive (39 in total; 12 *UNG*, 5 *NEIL1* and 22 *POLβ)*. There are 2 insertions (chr12:109,541,435 and chr12:109,541,462) and 10 SNPs (rs79744525, chr12:109,545,819, chr12:109,545,820, chr12:109,548,571, chr12:109,541,004, chr12:109,541,007, chr12:109,541,008, chr12:109,541,009, chr12:109,541,043 and chr12:109,541,051) among the false positive *UNG* gene variants. There are five false positive *NEIL1* gene variants with 1 insertion (chr15:75,645,566), 1 deletion (chr15:75,642,571) and 3 SNPs (chr15:75,645,207, chr15:75,645,209 and chr15:75,645,238). Most of the *POLβ* gene variants were found to be false positive. In 22 false positive results there are 17 SNPs (chr8:42,197,233, chr8:42,201,087, chr8:42,201,998, chr8:42,201,999, chr8:42,204,931, chr8:42,216,968, rs569312390, chr8:42,220,387, chr8:42,220,396, chr8:42,221,065, chr8:42,221,070, chr8:42,221,455, chr8:42,221,461, chr8:42,221,463, rs3136806, chr8:42,204,068, chr8:42,223,394) 2 insertions

(chr8:42,227,356 and chr8:42,204,551) and 3 deletions (chr8:42,201,087, chr8:42,216,970, chr8:42,219,877). Except *UNG* rs79744525 and *POLβ* rs3136806 gene variants all false positive results belong to novel variations. The IGV presentations with reading counts and Sanger sequencing chromatograms of the false positive results are shown in Figure 4.4.

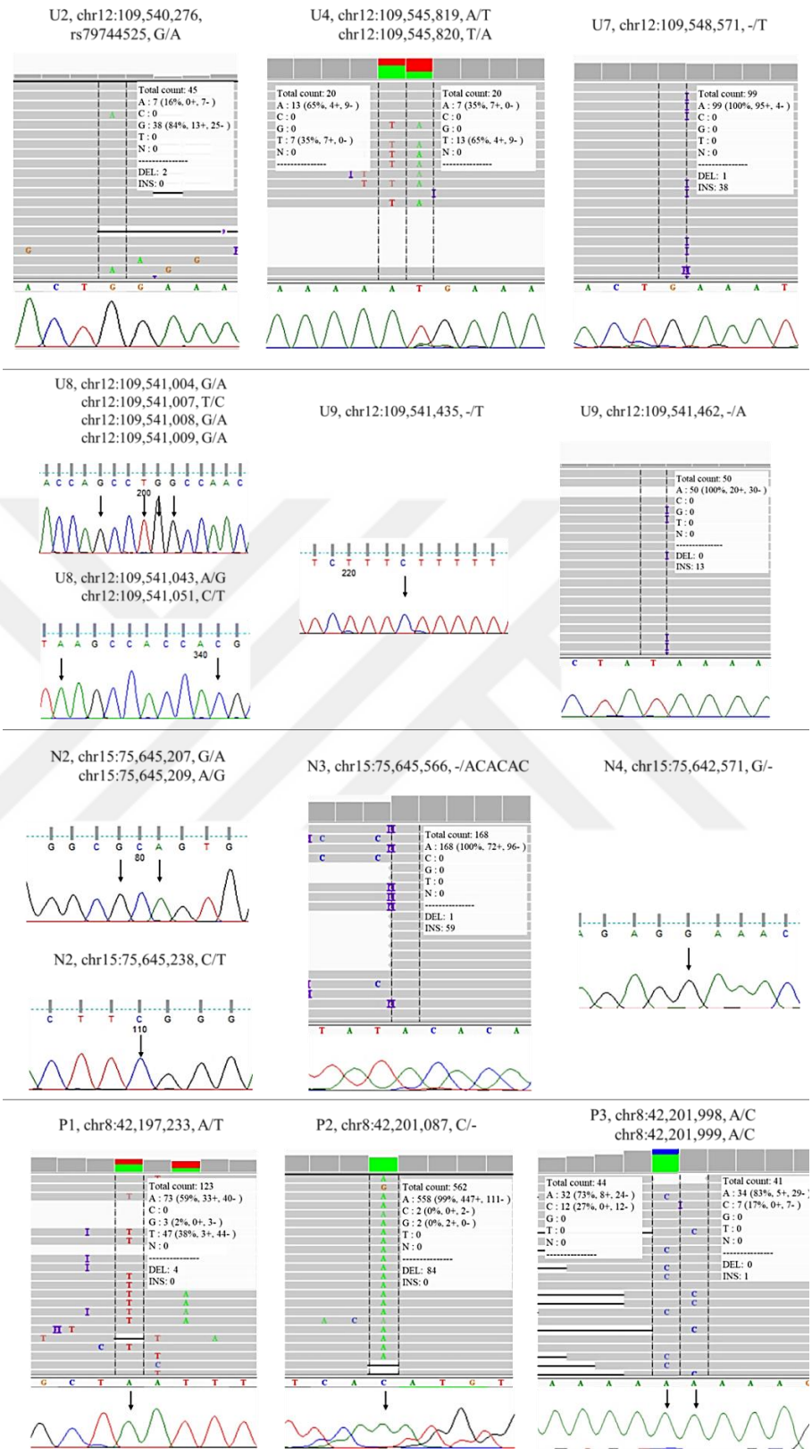

Figure 4.4. False positive Sanger sequencing results. IGV presentations are aligned with Sanger sequencing chromatograms. The reading counts of the variants are also included U, *UNG*; N, *NEIL1*; P, *POLβ*; A, *APOE*.

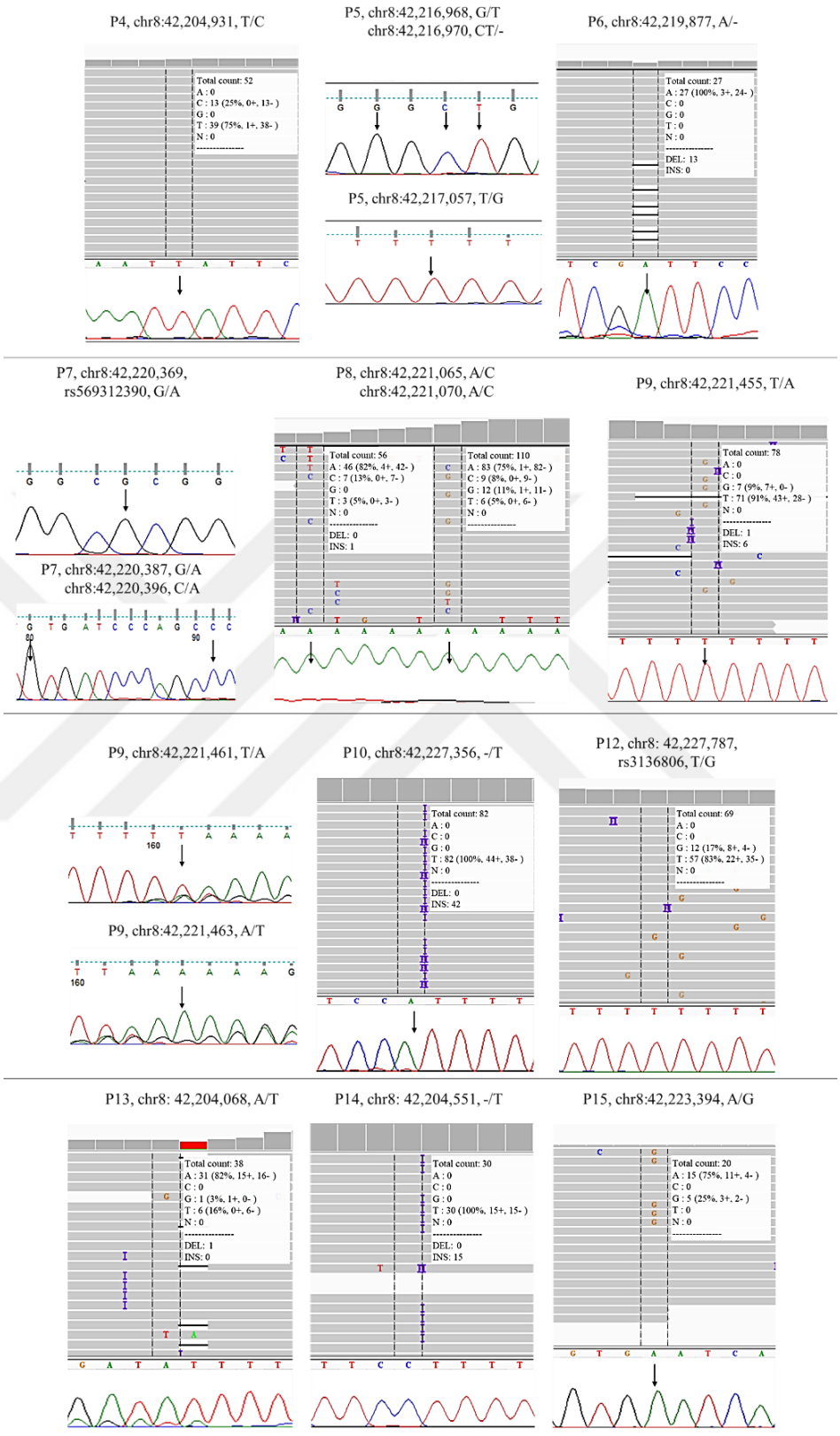

Figure 4.5. False positive Sanger sequencing results (continued). IGV presentations are aligned with Sanger sequencing chromatograms. The reading counts of the variants are also included U, *UNG*; N, *NEIL1*; P, *POLβ*; A, *APOE*.

4.5. Statistical Analyses of Blood Samples

Statistically significantly related LOAD-blood gene variants and their age matched control's allele and genotype frequencies are presented in Table 4.3. For better understanding the impact of the variants on LOAD the ORs with 95% CIs and Pearson $\chi 2$ values were also calculated. All of the gene variants represented in table are in Hardy-Weinberg equilibrium. Four SNPs minor alleles; rs2268406 (G), rs80001089 (G), rs1018782 (G) and rs1018783 (G) and 1 insertion rs1610925 (TA) of *UNG* gene were found to be statistically significantly related to increased LOAD-blood risk ($p<0.05$, $OR>1$). The heterozygous changes of rs1610925, rs2268406 and rs80001089 were also statistically significantly related to increased LOAD-blood risk ($p<0.05$, $OR>1$). *UNG* gene variant rs2430678 was also worth noticing ($p=0.0575$). The *NEIL1* and *POLβ* gene variants did not show statistically significant relation to LOAD. *POLβ* intronic variants rs3136806 SNP ($p = 0.0683$), rs35609234 INDEL ($p = 0.0706$) and rs11990332 SNP ($p = 0.0850$) are worth noticing. One mir631 insertion rs10653888 ($p=0.1618$) and SNP rs767369942 ($p=0.2839$) were found with a higher frequency in LOAD-blood but showed no statistically significant relation when compared with controls.

4.5.1. Allele and genotype frequencies of sporadic PD-blood samples and comparison with LOAD-blood

The allele and genotype frequency of *APOE* gene variant rs769446 in sporadic PD-blood and age matched control group is presented in Table 4.4. Both C allele ($p=0.0150$) and T/C genotype ($p=0.0114$) was statistically significantly related with sporadic PD when compared with controls. BER gene variants (chr8:42,221,070, chr8:42,201,087, chr8:42,221,065 chr15:75,645,566), which could be considered as a risk factor for sporadic PD because they are statistically significantly related ($p<0.05$) was found as false positive with Sanger sequencing. The *APOE* gene variant rs769446 was found with 0.11 frequency in LOAD-blood and with 0.15 frequency in age matched control group, hence found with no statistical significance.

Table 4.3. Allele and genotype frequencies of *UNG, POLβ, NEIL1* and *APOE* in peripheral blood samples of LOAD patients and age-matched cognitively normal controls. Bold; p-value <0

	Allele Frequency					Genotype Frequency						
	Allele	**LOAD**	**CTRL**	**OR (95% CI)**	**Fisher`s p-value**	**Genotype**	**LOAD**	**CTRL**	**OR (95% CI)**	**Fisher`s p-value**	**Pearson $\chi 2$**	**P value**
***UNG*, chr12**												
rs1610925	**-**	**0.803**	**0.913**	**2.58 (1.48-4.50)**	**0.0005**	**-/-**	**0.662**	**0.827**			**8.772**	**0.0031**
	TA	**0.197**	**0.087**			**-/TA**	**0.283**	**0.173**	**2.04 (1.11-3.75)**	**0.0224**		
						TA/TA	**0.055**	**0.000**	**-**	**0.0082**		
						-/TA+TA/TA	**0.338**	**0.173**	**2.44 (1.34-4.44)**	**0.0038**		
rs2268406	**T**	**0.836**	**0.918**	**2.21 (1.24-3.93)**	**0.0052**	**T/T**	**0.692**	**0.837**			**7.143**	**0.0075**
	G	**0.164**	**0.082**			**T/G**	**0.288**	**0.163**	**2.13 (1.15-3.96)**	**0.0153**		
						G/G	**0.020**	**0.000**	**-**	**0.2993**		
						T/G+G/G	**0.308**	**0.163**	**2.28 (1.23-4.22)**	**0.0075**		
rs80001089	**T**	**0.891**	**0.954**	**2.53 (1.21-5.30)**	**0.0129**	**T/T**	**0.788**	**0.908**			**6.651**	**0.0099**
	G	**0.109**	**0.046**			**T/G**	**0.207**	**0.092**	**2.60 (1.21-5.60)**	**0.0131**		
						G/G	**0.005**	**0.000**	**-**	**1.0000**		
						T/G+G/G	**0.212**	**0.092**	**2.66 (1.24-5.72)**	**0.0091**		
rs1018782	**A**	**0.848**	**0.918**	**2.01 (1.12-3.59)**	**0.0184**	**A/A**	**0.717**	**0.837**			**5.091**	**0.0241**
	G	**0.152**	**0.082**			**A/G**	**0.263**	**0.163**	**1.88 (1.01-3.50)**	**0.0563**		
						G/G	**0.020**	**0.000**	**-**	**0.2995**		
						A/G+G/G	**0.283**	**0.163**	**2.02 (1.09-3.75)**	**0.0304**		
rs1018783	**T**	**0.866**	**0.923**	**1.86 (1.02-3.40)**	**0.0406**	**T/T**	**0.753**	**0.847**			**3.448**	**0.0633**
	A	**0.134**	**0.077**			**T/A**	**0.227**	**0.153**	**1.67 (0.88-3.18)**	**0.1271**		
						A/A	**0.020**	**0.000**	**-**	**0.3004**		
						T/A+A/A	**0.247**	**0.153**	**1.82 (0.96-3.44)**	**0.0723**		
rs2430678	A	0.980	1.000	-	0.0575	A/A	0.960	1.000			4.070	0.0437
	G	0.020	0.000			A/G	0.040	0.000	-	0.0558		
						G/G	0.000	0.000	-	1.0000		
						A/G+G/G	0.040	0.000	-	0.0558		

***NEIL1*, chr15**												
rs10653888	- ACACACAC	0.967 0.033	0.990 0.010	3.29 (0.74-14.74)	0.1618	-/- -/ACACACAC ACACACAC/ ACACACAC -/ACACACAC + ACACACAC/ ACACACAC	0.934 0.066 0.000 0.066	0.980 0.020 0.000 0.020	 3.37 (0.75-15.25) - 3.37 (0.75-15.25)	 0.1563 1.0000 0.1563	2.790	0.0949
rs5745916	G A	0.942 0.058	0.969 0.031	1.95 (0.78-4.88)	0.1619	G/G G/A A/A G/A+A/A	0.884 0.116 0.000 0.116	0.939 0.061 0.000 0.061	 2.01 (0.79-5.12) - 2.01 (0.79-5.12)	 0.1511 1.0000 0.1511	2.239	0.1346
rs11634109	T C	0.891 0.109	0.923 0.077	1.47 (0.80-2.72)	0.2422	T/T T/C C/C T/C+C/C	0.788 0.207 0.005 0.212	0.847 0.153 0.000 0.153	 1.45 (0.76-2.78) - 1.49 (0.78-2.85)	 0.2746 1.0000 0.2735	1.471	0.2252
rs767369942	G A	0.980 0.020	0.995 0.005	4.02 (0.50-32.38)	0.2839	G/G G/A A/A G/A+A/A	0.960 0.040 0.000 0.040	0.990 0.010 0.000 0.010	 4.08 (0.50-33.13) - 4.08 (0.50-33.13)	 0.2801 1.0000 0.2801	2.028	0.1544
***POLβ*, chr8**												
rs3136806	T G	0.912 0.088	0.954 0.046	2.01 (0.95-4.28)	0.0683	T/T T/G G/G T/G+G/G	0.823 0.177 0.000 0.177	0.908 0.092 0.000 0.092	 2.12 (0.98-4.62) - 2.12 (0.98-4.62)	 0.0575 1.0000 0.0575	3.736	0.0532
rs35609234	C -	0.871 0.129	0.923 0.077	1.78 (0.97-3.26)	0.0706	C/C C/- -/- C/-+-/-	0.763 0.217 0.020 0.237	0.857 0.133 0.010 0.143	 1.84 (0.94-3.62) 2.22 (0.24-20.23) 1.87 (0.97-3.59)	 0.0836 0.6585 0.0672	3.579	0.0585
rs11990332	A G	0.881 0.119	0.929 0.071	1.75 (0.94-3.26)	0.0850	A/A A/G	0.783 0.197	0.867 0.122	 1.78 (0.89-3.58)	 0.1403	3.053	0.0806

						G/G	0.020	0.011	2.19 (0.24-19.94)	0.6597		
						A/G+G/G	0.217	0.133	1.81 (0.92-3.56)	0.0852		
rs571459229	G	0.985	1.000	-	0.1853	G/G	0.970	1.000			3.031	0.0817
	T	0.015	0.000			G/T	0.030	0.000	-	0.1830		
						T/T	0.000	0.000	-	1.0000		
						G/T+T/T	0.030	0.000	-	0.1830		
rs3136744	A	0.957	0.974	1.71 (0.62-4.71)	0.3608	A/A	0.914	0.949			1.156	0.2822
	C	0.043	0.026			A/C	0.086	0.051	1.75 (0.62-4.88)	0.3513		
						C/C	0.000	0.000	-	1.0000		
						A/C+C/C	0.086	0.051	1.75 (0.62-4.88)	0.3513		
***APOE*, chr19**												
rs769449	**G**	**0.866**	**0.974**	**5.90 (2.32-15.02)**	**0.0001**	**G/G**	**0.747**	**0.949**			**17.594**	**0.0001**
	A	**0.134**	**0.026**			**G/A**	**0.237**	**0.051**	**5.91 (2.27-15.39)**	**0.0001**		
						A/A	**0.015**	**0.000**	**-**	**0.2895**		
						G/A+A/A	**0.253**	**0.051**	**6.28 (2.42-16.33)**	**0.0001**		
rs429358	**T**	**0.811**	**0.939**	**3.58 (1.89-6.76)**	**0.0001**	**T/T**	**0.652**	**0.878**			**16.851**	**0.0001**
	C	**0.189**	**0.061**			**T/C**	**0.318**	**0.122**	**3.50 (1.78-6.87)**	**0.0001**		
						C/C	**0.030**	**0.000**	**-**	**0.0839**		
						T/C+C/C	**0.348**	**0.133**	**3.83 (1.96-7.50)**	**0.0001**		

Table 4.4. Allele and genotype frequencies of *UNG, POLβ, NEIL1* and *APOE* in peripheral blood samples of sporadic PD patients and age-matched cognitively normal controls. Bold; p-value <0

Allele Frequency						**Genotype Frequency**						
	Allele	**Sporadic PD**	**CTRL**	**OR (95% CI)**	**Fisher`s p-value**	**Genotype**	**Sporadic PD**	**CTRL**	**OR (95% CI)**	**Fisher`s p-value**	**Pearson $\chi 2$**	**P value**
***APOE*, chr19**												
rs769446	**T**	**0.875**	**0.970**	**5.22 (1.37-19.90)**	**0.0150**	**T/T**	**0.750**	**0.940**			**5.433**	**0.0198**
	C	**0.125**	**0.030**			**T/C**	**0.250**	**0.060**	**4.62 (1.26-16.92)**	**0.0114**		
						C/C	**0.000**	**0.000**	**-**	**-**		
						T/C+C/C	**0.250**	**0.060**	**4.62 (1.26-16.92)**	**0.0114**		

4.5.2. Minor allele frequencies of LOAD-blood statistically significant variations

Minor allele frequencies (MAFs) of the statistically significant variants in LOAD-blood Turkish population were summarized in Table 4.5. The MAFs for the Turkish population were calculated for the minor alleles (MAs) using the allele counts. In general the MAFs were correlated with 1000 Genome Project Phase 3 study populations MAFs; all 1000G phase 3 individuals (ALL); American (AMR), East Asian (EAS), European (EUR) and South Asian (SAS) except for the African (AFR) population. There seems to be an inverse proportion between all populations including Turkish population and AFR.

Table 4.5. Statistically significant *UNG* and *APOE* variant's MAFs.

Variations	MA	MAF in our population	ALL	AFR	AMR	EAS	EUR	SAS
rs1610925	TA	0.12	0.25	0.42	0.18	0.10	0.22	0.23
rs2268406	G	0.14	0.14	0.22	0.12	0.04	0.16	0.15
rs80001089	G	0.09	0.03	0.00	0.05	0.00	0.09	0.03
rs1018782	G	0.13	0.21	0.44	0.15	0.04	0.16	0.15
rs1018783	A	0.12	0.19	0.40	0.15	0.04	0.16	0.15
rs2430678	A	0.01	0.05	0.03	0.03	0.06	0.04	0.07
rs769449	G	0.10	0.06	0.01	0.08	0.08	0.12	0.06
rs429358	C	0.15	0.15	0.27	0.10	0.09	0.16	0.09

4.5.3. *APOE* allele and genotype frequencies and comparison with BER genes in LOAD-blood samples

Allele and genotype frequencies of *APOE* ε2, ε3 and ε4 alleles and ε2/ε2, ε2/ε3, ε2/ε4, ε3/ε3, ε3/ε4 and ε4/ε4 in blood samples of LOAD and CTRL are shown in Table 4.6. ε4 allele frequency (p=0.0001, OR= 3.58, 95% CI = 1.89-6.77) and ε3/ε4 (p=0.0001, OR= 4.07, 95% CI = 1.97-8.42) genotype frequency was found significiantly higher in LOAD-blood samples. Greatest risk factor of *APOE*, ε4/ε4 genotype frequency was also higher in LOAD blood samples but showed no significant interaction (p=0.0834).

Table 4.6. Allele and genotype frequencies of *APOE* ε2, ε3 and ε4 in blood samples. Bold; p-value <0

rs429358-rs7412					
		LOAD Frequency	**Control Frequency**	**OR (95% CI)**	**Fisher's p-value**
Allele					
ε2	T-T	0.0354	0.0408	1.00 (0.41-2.44)	1.0000
ε3	T-C	0.7753	0.8980	1 (REF)	REF
ε4	**C-C**	**0.1894**	**0.0612**	**3.58 (1.89-6.77)**	**0.0001**
Genotype					
ε2/ε2	TT-TT	0.000	0.000	-	1.0000
ε2/ε3	TT-TC	0.056	0.061	1.24 (0.44-3.50)	0.7993
ε2/ε4	TC-TC	0.015	0.020	1.02 (0.17-6.22)	1.0000
ε3/ε3	TT-CC	0.601	0.816	1 (REF)	REF
ε3/ε4	**TC-CC**	**0.298**	**0.102**	**4.07 (1.97-8.42)**	**0.0001**
ε4/ε4	CC-CC	0.030	0.000	-	0.0834

The relation between statistically significantly related variants found LOAD-blood and *APOE* ε4 carrying status was investigated. During the calculation of ε4 allele frequency ε2/ε4 genotype carriers (3 LOAD and 2 controls) were excluded due to the ε2 allele's protective effect for LOAD. First of all the effect of statistically significantly related variants on *APOE* ε4 carrying status was investigated. For the *APOE* ε4 carriers *APOE* rs769449 variation contributed to the LOAD risk. On the other hand *UNG* rs1610925, rs2268406, rs80001089, rs1018782 and rs1018783 showed an additive effect to LOAD in *APOE* ε4 non-carriers. *UNG* rs2430678 variant showed no effect to LOAD phenotype in *APOE* ε4 non-carriers. Also, effect of *APOE* ε4 carrying status on statistically significantly related variants LOAD relation was investigated. *APOE* ε4 carrying showed an additive effect to LOAD in all statistically significantly related *UNG* variants non carriers (rs1610925, rs2268406, rs80001089, rs1018782, rs1018783 and rs2430678). *APOE* ε4 carrying status showed no significant effect on *APOE* rs769449 variation. Table 4.7 summarizes the relation between statistically significantly related variants and *APOE* ε4 carrying status.

Table 4.7. Statistically significantly related variants and APOE ε4 carrying status interaction in LOAD-blood samples. Bold, p-value <0; +, carrier; -, non-carrier

		LOAD	**Control**	**OR (95% CI)**	**Fisher's p-value**
APOE ε4 (+)	**rs769449 (+)**	**24%**	**3%**	**6.09 (1.42-26.17)**	**0.0131**

APOE ε4 (-)	rs769449 (-)	9%	7%		
	rs769449 (+)	1%	0%	-	1.0000
	rs769449 (-)	66%	90%		
APOE ε4 (+)	rs1610925 (+)	11%	4%	0.70 (0.18-2.75)	0.7210
	rs1610925 (-)	23%	6%		
APOE ε4 (-)	**rs1610925 (+)**	**24%**	**11%**	**4.00 (1.94-8.26)**	**0.0001**
	rs1610925 (-)	42%	79%		
APOE ε4 (+)	rs2268406 (+)	10%	4%	0.67 (0.17-2.62)	0.7173
	rs2268406 (-)	23%	6%		
APOE ε4 (-)	**rs2268406 (+)**	**24%**	**11%**	**3.83 (1.85-7.93)**	**0.0001**
	rs2268406 (-)	43%	79%		
APOE ε4 (+)	rs80001089 (+)	6%	3%	0.53 (0.12-2.35)	0.4084
	rs80001089 (-)	27%	7%		
APOE ε4 (-)	**rs80001089 (+)**	**15%**	**4%**	**6.03 (2.04-17.84)**	**0.0002**
	rs80001089 (-)	52%	86%		
APOE ε4 (+)	rs1018782 (+)	8%	3%	0.70 (0.16-3.05)	0.6950
	rs1018782 (-)	26%	7%		
APOE ε4 (-)	**rs1018782 (+)**	**21%**	**11%**	**3.11 (1.50-6.48)**	**0.0019**
	rs1018782 (-)	46%	79%		
APOE ε4 (+)	rs1018783 (+)	6%	3%	0.53 (0.12-2.35)	0.4084
	rs1018783 (-)	27%	7%		
APOE ε4 (-)	**rs1018783 (+)**	**19%**	**11%**	**2.78 (1.33-5.82)**	**0.0073**
	rs1018783 (-)	48%	79%		
APOE ε4 (+)	rs2430678 (+)	2%	0%	-	1.0000
	rs2430678 (-)	31%	10%		
APOE ε4 (-)	rs2430678 (+)	3%	0%	-	0.1600
	rs2430678 (-)	64%	90%		

rs769449 (+)	APOE ε4 (+)	24%	3%	-	1.0000
	APOE ε4 (-)	1%	0%		
rs769449 (-)	APOE ε4 (+)	9%	7%	1.75 (0.70-4.38)	0.2817
	APOE ε4 (-)	66%	90%		
rs1610925 (+)	APOE ε4 (+)	11%	4%	1.20 (0.34-4.22)	1.0000
	APOE ε4 (-)	24%	11%		
rs1610925 (-)	**APOE ε4 (+)**	**23%**	**6%**	**6.88 (2.78-17.02)**	**0.0001**
	APOE ε4 (-)	42%	79%		
rs2268406 (+)	APOE ε4 (+)	10%	4%	1.20 (0.34-4.21)	1.0000
	APOE ε4 (-)	24%	11%		
rs2268406 (-)	**APOE ε4 (+)**	**23%**	**6%**	**6.88 (2.78-17.02)**	**0.0001**
	APOE ε4 (-)	43%	79%		

		LOAD	**Control**	**OR (95% CI)**	**Fisher's p-value**
rs80001089 (+)	APOE ε4 (+)	6%	3%	0.55 (0.11-2.85)	0.6615
	APOE ε4 (-)	15%	4%		
rs80001089 (-)	**APOE ε4 (+)**	**27%**	**7%**	**6.30 (2.72-14.58)**	**0.0001**
	APOE ε4 (-)	52%	86%		
rs1018782 (+)	APOE ε4 (+)	8%	3%	1.38 (0.34-5.62)	0.7471

rs1018782 (-)	APOE ε4 (-)	21%	11%		
	APOE ε4 (+)	**26%**	**7%**	**6.11 (2.62-14.26)**	**0.0001**
	APOE ε4 (-)	46%	79%		
rs1018783 (+)	APOE ε4 (+)	6%	3%	1.19 (0.28-4.98)	1.0000
	APOE ε4 (-)	19%	11%		
rs1018783 (-)	**APOE ε4 (+)**	**27%**	**7%**	**6.27 (2.70-14.58)**	**0.0001**
	APOE ε4 (-)	48%	79%		
rs2430678 (+)	APOE ε4 (+)	2%	0%	-	1.0000
	APOE ε4 (-)	3%	0%		
rs2430678 (-)	**APOE ε4 (+)**	**31%**	**10%**	**4.30 (2.09-8.83)**	**0.0001**
	APOE ε4 (+)	64%	90%		

4.5.4. Genetic models of statistically significant LOAD-blood variants

The statistically significantly related *UNG* variants were also according to additive, dominant and recessive genetic models. *UNG* rs1610925, rs2268406, rs1018782, rs1018783, rs2430678 variants were fitted better to additive model while rs80001089 variant was fitted better to the dominant model (Table 4.8).

Table 4.8. Genetic models of statistically significantly related *UNG* variants in LOAD-blood samples. Bold, p-value <0

	Additive Model		**Dominant Model**		**Recessive Model**	
	OR (95% CI)	**Fisher's p-value**	**OR (95% CI)**	**Fisher's p-value**	**OR (95% CI)**	**Fisher's p-value**
rs1610925	**2.58 (1.48–4.50)**	**0.0005**	**2.44 (1.34–4.44)**	**0.0038**	**-**	**0.0183**
rs2268406	**2.76 (1.56–4.87)**	**0.0002**	**2.62 (1.42–4.83)**	**0.0015**	**-**	**0.0183**
rs80001089	**2.53 (1.21–5.30)**	**0.0129**	**2.66 (1.24–5.72)**	**0.0091**	-	1.0000
rs1018782	**2.19 (1.23–3.90)**	**0.0184**	**2.02 (1.08–3.75)**	**0.0304**	-	0.3058
rs1018783	**1.86 (1.02–3.40)**	**0.0406**	1.82 (0.96–3.44)	0.0723	-	0.3058
rs2430678	-	0.0575	-	0.0558	-	1.0000

4.5.6. Haplotype analysis of statistically significant LOAD-blood samples

The linkage disequilibrium (LD) between statistically significantly related variants of LOAD-blood samples were studied using haplotype analysis. The measure of the level of recombination between blocks is represented with D' value (pairwise SNP correlation) while LOD means logarithm of odds. The LD of variants was represented in shades of red according to their power (Figure 4.6). The bright red boxes meaning is LD with D'=1 and LOD ≥ 2 while white boxes meaning is D' < 1 and LOD < 2. The LD distribution was grouped as blocks 1 to 4 according to the genes in the following order; *POLβ, APOE, UNG, NEIL1.* The following *UNG* variants were in complete LD;

rs1018782-rs1018783, rs1018783-rs1610925, rs1610925-rs80001089, rs1018782-rs1610925, rs1610925-rs2430687, rs1018782-rs2268406, rs1018783-rs2268406, rs1610925-rs2268406 and rs80001089-rs2268406. *APOE* variant pair rs769449-rs429358 was in strong LD and also it showed strong LD with *UNG* rs1018782-rs80001089 $r^2 \geq 0.50$ and D' approaching to 1. The variant pairs belonging to *NEIL1* and *POLβ* showed no strong LD (Figure 4.6).

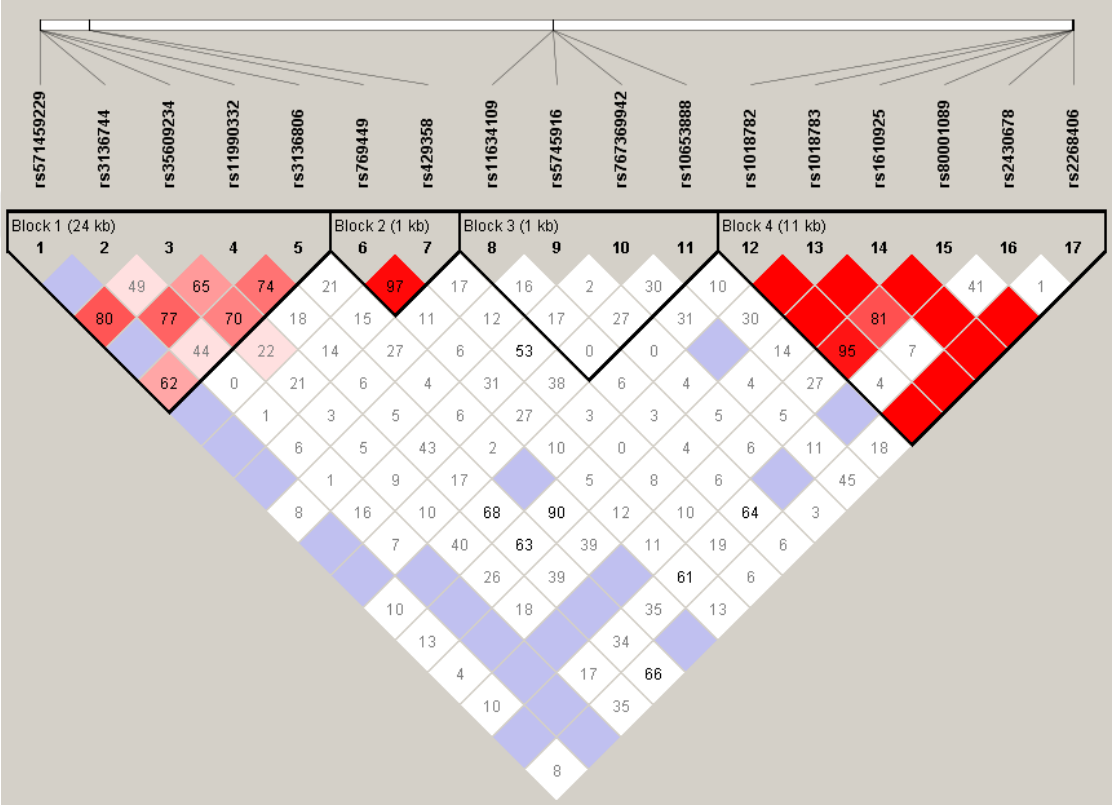

Figure 4.6. Linkage disequilibrium plot of LOAD. The blocks of the plot; 1, 2, 3 and 4 were defined according to the genes in the following order; *POLβ*, *APOE*, *NEIL1* and *UNG*. Bright red means strong LD while white means weak LD.

POLβ GACAT, *APOE* GT, *UNG* ATTTAT and *NEIL1* TGGA haplotypes were found in a significantly protective relation with LOAD (OR<1 and p<0.05). While *POLβ* GAAGG, *APOE* AC and *UNG* GAAGAG haplotype frequencies were statistically significantly higher in LOAD-blood cases (OR>1 and p<0.05) but odds ratio of *POLβ* GAAGG could not be calculated because of its absence in controls. (n=0). No *NEIL1* haplotype was found in a significant relation with LOAD-blood cases. In table 4.9, haplotypes were grouped according to their blocks, thus their genes and the frequencies, OR and Chi Square calculations with p values were summarized.

Table 4.9. Haplotype analysis of statistically significant variants. The blocks were created according to the genes. Bold, p-value <0

Gene	SNP #	Haplotype	All (%)	Case (%)	Control (%)	OR (95% CI)	Fisher`s p-value	Chi Square	P Value
Block 1									
POLβ	1,2,3,4,5	**GACAT**	**0.902**	**0.871**	**0.965**	**0.24 (0.11-0.54)**	**0.0001**	**13.188**	**0.0003**
		GAAGG	0.021	0.032	0.000	-	0.0063	6.494	0.0108
Block 2									
APOE	6,7	**GT**	**0.853**	**0.810**	**0.939**	**0.27 (0.15-0.52)**	**0.0001**	**17.529**	**0.00003**
		AC	**0.096**	**0.131**	**0.025**	**5.83 (2.29-14.86)**	**0.0001**	**17.101**	**0.00004**
		GC	0.049	0.056	0.035	1.60 (0.67-3.82)	0.3194	1.168	0.2798
Block 3									
NEIL1	8,9,10,11	**TGGA**	**0.837**	**0.813**	**0.886**	**0.57 (0.34-0.94)**	**0.0336**	**5.185**	**0.0228**
		CGGA	0.089	0.097	0.073	0.49 (0.30-0.79)	0.2888	0.914	0.339
		TAGA	0.035	0.040	0.023	2.04 (0.67-6.19)	0.2354	1.222	0.269
		TGGC	0.012	0.013	0.010	1.25 (0.24-6.52)	1.0000	0.073	0.7866
Block 4									
UNG	12,13,14, 15,16,17	**ATTTAT**	**0.848**	**0.816**	**0.914**	**0.42 (0.24-0.73)**	**0.0015**	**9.959**	**0.0016**
		GAAGAG	**0.074**	**0.091**	**0.040**	**2.37 (1.08-5.21)**	**0.0205**	**4.909**	**0.0267**
		GAATAG	0.040	0.043	0.035	1.22 (0.50-3.00)	0.8258	0.195	0.6585
		ATATGT	0.012	0.018	0.000	-	0.1018	2.031	0.1541
		GTAGAG	0.010	0.013	0.005	2.52 (0.29-21.71)	0.6691	0.758	0.3841

4.5.7. Epistatic relations of statistically significant LOAD-blood samples

For better understanding the combinatorial effects of LOAD-blood significant variants, epistatic relation analysis was done. The epistatic relation between rs80001089-rs1610925; rs1610925-rs2268406; rs80001089-rs2268406; rs80001089-rs1018782; rs2268406-rs1018782; rs1610925-rs1018782 *UNG* variants and rs429358-rs769449 *APOE* variants were found statistically significant (p<0.05). No *NEIL1* or *POLβ* gene variants or variants between different genes found in significant epistatic relation (Table 4.10).

Table 4.10. Epistatic relation of LOAD-blood significant variants. Frequencies in LOAD and Control group with their OR and p values were showed. Bold, p-value <0

	Variations	LOAD %	Control %	OR (95% CI)	Fisher`s p-value
APOE-APOE	**rs429358-rs769449**	**0.247**	**0.051**	**6.12 (2.35-15.91)**	**0.0001**

UNG-UNG	**rs80001089-rs1610925**	**0.212**	**0.092**	**2.66 (1.24-5.72)**	**0.0002**
UNG-UNG	**rs1610925-rs2268406**	**0.308**	**0.163**	**2.28 (1.23-4.22)**	**0.0075**
UNG-UNG	**rs80001089-rs2268406**	**0.212**	**0.092**	**2.66 (1.24-5.72)**	**0.0091**
UNG-UNG	**rs80001089-rs1018782**	**0.202**	**0.092**	**2.50 (1.16-5.40)**	**0.0194**
UNG-UNG	**rs2268406-rs1018782**	**0.283**	**0.163**	**2.02 (1.09-3.75)**	**0.0304**
UNG-UNG	**rs1610925-rs1018782**	**0.283**	**0.163**	**2.02 (1.09-3.75)**	**0.0304**
UNG-UNG	rs2268406-rs1018783	0.247	0.153	1.82 (0.96-3.44)	0.0723
UNG-UNG	rs1018782-rs1018783	0.247	0.153	1.82 (0.96-3.44)	0.0723
UNG-UNG	rs1018783-rs1610925	0.247	0.153	1.82 (0.96-3.44)	0.0723
APOE-UNG	rs769449-rs2268406	0.086	0.031	2.97 (0.85-10.41)	0.0876
APOE-UNG	rs769449-rs1610925	0.091	0.031	3.16 (0.91-11.02)	0.0894
APOE-UNG	rs769449-rs1018782	0.076	0.031	2.60 (0.73-9.20)	0.1948
POLβ – POLβ	rs11993638-rs3136788	0.081	0.041	2.06 (0.67-6.35)	0.2284
POLβ - POLβ	rs3136744-rs3136788	0.081	0.041	2.06 (0.67-6.35)	0.2284
APOE-UNG	rs429358-rs2268406	0.096	0.051	1.97 (0.71-5.46)	0.2578
APOE-NEIL1	rs769449-rs11634109	0.066	0.031	2.22 (0.62-7.99)	0.2795

4.5.8. Power analysis of statistically significant LOAD-blood variations

The power of the statistically significantly related variants was calculated using t-test and assuming two-tailed p value. All statistically significantly related variants have more than 80% power. *UNG* gene variant rs1610925 and *APOE* gene variant rs769449 has 85.6% power and also *UNG* gene variants rs80001089 and rs1018783 share the same power; %92.3. Statistically significantly related UNG gene variants rs2268406, rs1018782 and APOE gene variant rs429358 powers are; 90.1%, 91.3% and 81.8%, respectively.

4.6. Statistical Analyses of Post-mortem Brain Tissue Samples

All CE LOAD, hpC and control samples (11, 11 and 10 in order) were included for the calculations of allele and genotype frequencies, their ORs with 95% CIs and Pearson $\chi 2$ values in Table 4.11. *UNG* gene variant rs2569987 C allele was found statistically significantly related with LOAD when compared with control (p=0.0216). hpC individuals do not show LOAD symptoms yet have histopathological markers of LOAD because of that calculations were made also by including hpC individuals to control group. *UNG* gene variant rs2569987 were still statistically related with LOAD with an increased significance (p=0.0162). There was no statistically significantly related NEIL1 or *POLβ* gene variant and also *APOE* rs429358 did not show significant relation to LOAD in post-mortem CE tissues. When hpC and control samples were

compared, *APOE* gene variant rs405509 G allele was found statistically significantly related with hpC group.

Table 4.11. Allele and genotype frequencies of *UNG, POLβ, NEIL1* and *APOE* in CE samples of LOAD patients, age-matched cognitively normal and hpC subjects (LOAD=11, hpC=11, Control=10). Bold, p-value <0

LOAD vs Control												
	Allele frequency					**Genotype frequency**						
	Allele	**LOAD**	**Control**	**OR (95% CI)**	**Fisher's p-value**	**Genotype**	**LOAD**	**Control**	**OR (95% CI)**	**Fisher's p-value**	**Pearson χ2**	**P value**
***UNG*, chr12**												
rs2569987	**T**	**0.73**	**1.00**	**-**	**0.0216**	**T/T**	**0.55**	**1.00**			**5.966**	**0.0146**
	C	**0.27**	**0.00**			**T/C**	**0.36**	**0.00**	**-**	**0.0867**		
						C/C	**0.09**	**0.00**	**-**	**0.4118**		
						T/C+C/C	**0.45**	**0.00**	**-**	**0.0351**		
rs2268406	T	0.82	0.95	4.22 (0.43-41.45)	0.3465	T/T	0.64	0.90			2.007	0.1566
	G	0.18	0.05			T/G	0.36	0.10	5.14 (0.46-56.89)	0.3108		
						G/G	0.00	0.00	-	1.0000		
						T/G+G/G	0.36	0.10	5.14 (0.46-56.89)	0.3108		
rs80001089	T	0.91	1.00	-	0.4890	T/T	0.82	1.00			2.010	0.1563
	G	0.09	0.00			T/G	0.18	0.00	-	0.4762		
						G/G	0.00	0.00	-	1.0000		
						T/G+G/G	0.18	0.00	-	0.4762		
***NEIL1*, chr15**												
rs7182283	G	0.45	0.60	1.80 (0.53-6.14)	0.3743	G/G	0.28	0.50			1.147	0.2841
	T	0.55	0.40			G/T	0.36	0.20	3.33 (0.36-30.70)	0.5921		
						T/T	0.36	0.30	2.22 (0.28-17.63)	0.6193		
						G/T+T/T	0.72	0.50	2.67 (0.43-16.39)	0.3870		
***APOE*, chr19**												
rs429358	T	0.68	0.80	1.87 (0.45-7.69)	0.4913	T/T	0.45	0.60			0.444	0.5051
	C	0.32	0.20			T/C	0.46	0.40	1.50 (0.26-8.82)	1.0000		
						C/C	0.09	0.00	-	1.0000		
						T/C+C/C	0.55	0.40	1.80 (0.32-10.20)	0.6699		

hpC vs Control												
	Allele frequency					Genotype frequency						
	Allele	**hpC**	**Control**	**OR (95% CI)**	**Fisher's p-value**	**Genotype**	**LOAD**	**Control**	**OR (95% CI)**	**Fisher's p-value**	**Pearson $\chi 2$**	**P value**
***UNG*, chr12**												
rs2569987	T	0.91	1.00	-	0.4890	T/T	0.82	1.00			2.010	0.1563
	C	0.09	0.00			T/C	0.18	0.00	-	0.4762		
						C/C	0.00	0.00	-	1.0000		
						T/C+C/C	0.18	0.00	-	0.4762		
***POLβ*, chr8**												
rs4526344	G	0.86	1.00	-	0.2334	G/G	0.73	1.00			3.182	0.0745
	C	0.14	0.00			G/C	0.27	0.00	-	0.2143		
						C/C	0.00	0.00	-	1.0000		
						G/C+C/C	0.27	0.00	-	0.2143		
***APOE*, chr19**												
rs405509	**T**	**0.73**	**1.00**	**-**	**0.0216**	**T/T**	**0.55**	**1.00**			**5.966**	**0.0146**
	G	**0.27**	**0.00**			**T/G**	**0.36**	**0.00**	**-**	**0.0867**		
						G/G	**0.09**	**0.00**	**-**	**0.4118**		
						T/G+G/G	**0.45**	**0.00**	**-**	**0.0351**		
LOAD vs hpC+Control												
	Allele frequency					**Genotype frequency**						
	Allele	**LOAD**	**Control**	**OR (95% CI)**	**Fisher's p-value**	**Genotype**	**LOAD**	**Control**	**OR (95% CI)**	**Fisher's p-value**	**Pearson $\chi 2$**	**P value**
***UNG*, chr12**												
rs2569987	**T**	**0.73**	**0.95**	**7.50 (1.37-41.14)**	**0.0162**	**T/T**	**0.55**	**0.90**			**5.453**	**0.0195**
	C	**0.27**	**0.05**			**T/C**	**0.36**	**0.10**	**6.33 (0.92-43.62)**	**0.0674**		
						C/C	**0.09**	**0.00**	**-**	**0.2692**		
						T/C+C/C	**0.45**	**0.10**	**7.92 (1.21-51.84)**	**0.0318**		

rs2268406	T G	0.82 0.18	0.95 0.05	4.44 (0.75-26.52)	0.1697	T/T T/G G/G T/G+G/G	0.64 0.36 0.00 0.36	0.90 0.10 0.00 0.10	 5.43 (0.81-36.51) - 5.43 (0.81-36.51)	 0.1476 1.0000 0.1476	3.413	0.0647
rs80001089	T G	0.91 0.09	1.00 0.00	-	0.1146	T/T T/G G/G T/G+G/G	0.82 0.18 0.00 0.18	1.00 0.00 0.00 0.00	 - - -	 0.1109 1.0000 0.1109	4.073	0.0436
***NEIL1*, chr15**												
75,641,932	A G	0.91 0.09	1.00 0.00	-	0.1146	A/A A/G G/G A/G+G/G	0.82 0.18 0.00 0.18	1.00 0.00 0.00 0.00	 - - -	 0.1109 1.0000 0.1109	4.073	0.0436
rs7182283	G T	0.45 0.55	0.60 0.40	1.76 (0.62-4.99)	0.3044	G/G G/T T/T G/T+T/T	0.27 0.36 0.36 0.73	0.43 0.33 0.24 0.57	 1.71 (0.29-10.30) 2.40 (0.38-15.32) 2.00 (0.41-97.74)	 0.6668 0.3972 0.4647	0.748	0.3871
***POLβ*, chr8**												
rs3136791	T G	0.91 0.09	0.98 0.02	4.10 (0.35-47.96)	0.2698	T/T T/G G/G T/G+G/G	0.82 0.18 0.00 0.18	0.95 0.05 0.00 0.05	 4.44 (0.36-55.58) - 4.44 (0.36-55.58)	 0.2661 1.0000 0.2661	1.800	0.1798
rs2976238	T G	0.68 0.32	0.81 0.19	1.98 (0.61-6.47)	0.3522	T/T T/G G/G T/G+G/G	0.45 0.45 0.09 0.55	0.62 0.38 0.00 0.38	 1.63 (0.36-7.43) - 1.95 (4.44-8.55)	 0.7007 0.3158 0.4651	0.794	0.3730
***APOE*, chr19**												
rs429358	T C	0.68 0.32	0.83 0.17	2.33 (0.70-7.82)	0.2079	T/T T/C C/C T/C+C/C	0.45 0.45 0.09 0.55	0.67 0.33 0.00 0.33	 2.00 (0.43-9.29) - 2.40 (0.54-10.69)	 0.4472 0.3000 0.2826	1.347	0.2459

Targeted NGS analyses of both brain tissues (CE and TC) belonging to patients AD4, CT2, AS1, AS2 and AS5 could not be completed due to TC sample's low DNA quality. For comparing the similarities and differences between different brain regions, the allele and genotype frequencies of the same patients' CE and TC samples (10 LOAD, 9 cognitively normal controls and 8 hpC) are presented in Table 4.12.

In CE samples, when the sample size decreased, *UNG* gene variant rs2569987 did not show statistical significance (p=0.1071) when compared with, only control and control+hpC, because AD4 was rs2569987 carrier. APOE gene variant rs405509 remained statistically significantly related with hpC when compared with control (p=0.0060). Still *NEIL1* and *POLβ* gene variants showed no statistically significant relation with LOAD or hpC in CE (Table 4.12).

In TC samples, *POLβ* gene variant rs1012381950 T/C genotype was found statistically significantly related (p=0.0198) with LOAD-TC when compared with Control-TC (Table 4.12). *UNG* gene variant rs80001089 G allele (p=0.0153) and T/G genotype (p=0.0120) was found statistically significantly related with LOAD in comparison with hpC+Control. But when LOAD compared with controls only the significance was lost.

The allele (p = 0.0153) and genotype (p = 0.012) frequency of *UNG* rs80001089 were found to be statistically significantly different between TC-LOAD and TC-Control+hpC (Table 4.12). But, p=1071 for *UNG* rs80001089 when comparing LOAD-TC and control-TC (Table 4.12).

Table 4.12. Allele and genotype frequencies of *UNG, POLβ, NEIL1* and *APOE* in CE and TC samples of same LOAD patients, age-matched cognitively normal and hpC subjects (LOAD=10, hpC=8, Control=9). Bold, p-value <0.

LOAD vs Control, CE												
	Allele frequency					**Genotype frequency**						
	Allele	**LOAD**	**Control**	**OR (95% CI)**	**Fisher's p-value**	**Genotype**	**LOAD**	**Control**	**OR (95% CI)**	**Fisher's p-value**	**Pearson χ2**	**P value**
***UNG*, chr12**												
rs2569987	T	0.80	1.00	-	0.1071	T/T	0.60	1.00			4.560	0.0327
	C	0.20	0.00			T/C	0.40	0.00	-	0.0867		
						C/C	0.00	0.00	-	1.0000		
						T/C+C/C	0.40	0.00	-	0.0867		
rs2268406	T	0.80	0.94	4.25 (0.43-42.19)	0.3436	T/T	0.60	0.89			2.039	0.1533
	G	0.20	0.06			T/G	0.40	0.11	5.33 (0.47-60.80)	0.3034		
						G/G	0.00	0.00	-	1.0000		
						T/G+G/G	0.40	0.11	5.33 (0.47-60.80)	0.3034		
rs80001089	T	0.90	1.00	-	0.4879	T/T	0.80	1.00			2.012	0.1561
	G	0.10	0.00			T/G	0.20	0.00	-	0.4737		
						G/G	0.00	0.00	-	1.0000		
						T/G+G/G	0.20	0.00	-	0.4737		
***NEIL1*, chr15**												
rs7182283	G	0.40	0.67	3.00 (0.80-11.31)	0.1192	G/G	0.20	0.56			2.574	0.1087
	T	0.60	0.33			G/T	0.40	0.22	5.00 (0.47-52.96)	0.2865		
						T/T	0.40	0.22	5.00 (0.47-52.96)	0.1698		
						G/T+T/T	0.80	0.44	5.00 (0.66-38.15)	0.2865		
***APOE*, chr19**												
rs429358 Cys130Arg	T	0.65	0.78	1.88 (0.45-7.97)	0.4848	T/T	0.40	0.56			0.460	0.4977
	C	0.35	0.22			T/C	0.50	0.44	1.56 (0.24-10.03)	1.0000		
						C/C	0.10	0.00	-	1.0000		
						T/C+C/C	0.60	0.44	1.88 (0.30-11.63)	0.6563		

hpC vs Control, CE												
	Allele frequency					Genotype frequency						
	Allele	**hpC**	**Control**	**OR (95% CI)**	**Fisher's p-value**	**Genotype**	**LOAD**	**Control**	**OR (95% CI)**	**Fisher's p-value**	**Pearson $\chi 2$**	**P value**
***UNG*, chr12**												
rs2569987	T	0.88	1.00	-	0.2139	T/T	0.75	1.00			2.550	0.1103
	C	0.13	0.00			T/C	0.25	0.00	-	0.2059		
						C/C	0.00	0.00	-	1.0000		
						T/C+C/C	0.25	0.00	-	0.2059		
***POLβ*, chr8**												
rs4526344	G	0.88	1.00	-	0.2139	G/G	0.75	1.00			2.550	0.1103
	C	0.13	0.00			G/C	0.25	0.00	-	0.2059		
						C/C	0.00	0.00	-	1.0000		
						G/C+C/C	0.25	0.00	-	0.2059		
***APOE*, chr19**												
rs405509	**T**	**0.63**	**1.00**	**-**	**0.0060**	**T/T**	**0.38**	**1.00**			**7.969**	**0.0048**
	G	**0.37**	**0.00**			**T/G**	**0.50**	**0.00**	**-**	**0.0192**		
						G/G	**0.12**	**0.00**	**-**	**0.3077**		
						T/G+G/G	**0.62**	**0.00**	**-**	**0.0090**		
LOAD vs hpC+Control, CE												
	Allele frequency					**Genotype frequency**						
	Allele	**LOAD**	**Control**	**OR (95% CI)**	**Fisher's p-value**	**Genotype**	**LOAD**	**Control**	**OR (95% CI)**	**Fisher's p-value**	**Pearson $\chi 2$**	**P value**
***UNG*, chr12**												
rs2569987	T	0.80	0.94	4.00 (0.66-24.21)	0.1792	T/T	0.60	0.88			2.904	0.0883
	C	0.20	0.06			T/C	0.40	0.12	5.00 (0.72-34.94)	0.1535		
						C/C	0.00	0.00	-	1.0000		
						T/C+C/C	0.40	0.12	5.00 (0.72-34.94)	0.1535		

rs2268406	T G	0.80 0.20	0.94 0.06	4.00 (0.66-24.21)	0.1792	T/T T/G G/G T/G+G/G	0.60 0.40 0.00 0.40	0.88 0.12 0.00 0.12	 5.00 (0.72-34.94) - 5.00 (0.72-34.94)	 0.1535 1.0000 0.1535	2.904	0.0883
rs80001089	T G	0.90 0.10	1.00 0.00	-	0.1328	T/T T/G G/G T/G+G/G	0.80 0.20 0.00 0.20	1.00 0.00 0.00 0.00	 - - -	 0.1282 1.0000 0.1282	3.672	0.0553
***NEIL1*, chr15**												
75,641,932	A G	0.90 0.10	1.00 0.00	-	0.1328	A/A A/G G/G A/G+G/G	0.80 0.20 0.00 0.20	1.00 0.00 0.00 0.00	 - - -	 0.1282 1.0000 0.1282	3.672	0.0553
rs7182283	T C	0.40 0.60	0.59 0.41	2.14 (0.70-6.60)	0.2604	T/T T/C C/C T/C+C/C	0.20 0.40 0.40 0.80	0.41 0.35 0.24 0.59	 2.33 (0.31-17.55) 3.50 (0.43-28.45) 2.80 (0.45-17.38)	 0.6285 0.3348 0.4059	1.271	0.2597
***POLβ*, chr8**												
rs3136791	T G	0.90 0.10	0.97 0.03	3.67 (0.31-43.27)	0.5477	T/T T/G G/G T/G+G/G	0.80 0.20 0.00 0.20	0.94 0.06 0.00 0.06	 5.33 (0.41-70.20) - 5.33 (0.41-70.20)	 0.2313 1.0000 0.2313	1.883	0.1700
rs2976238	T G	0.65 0.35	0.76 0.24	1.75 (0.52-5.89)	0.5302	T/T T/G G/G T/G+G/G	0.40 0.50 0.10 0.60	0.53 0.47 0.00 0.47	 2.81 (0.42-18.74) - 3.38 (0.52-21.73)	 0.3864 0.2500 0.2337	1.724	0.1892
***APOE*, chr19**												
rs429358	T C	0.65 0.35	0.82 0.18	2.51 (0.70-8.98)	0.1936	T/T T/C C/C	0.40 0.50 0.10	0.65 0.35 0.00	 2.29 (0.44-11.92) -	 0.4185 0.3125	1.556	0.2122

						T/C+C/C	0.60	0.35	2.75 (0.55-13.75)	0.2566		
LOAD vs Control, TC												
	Allele frequency					**Genotype frequency**						
	Allele	**LOAD**	**Control**	**OR (95% CI)**	**Fisher's p-value**	**Genotype**	**LOAD**	**Control**	**OR (95% CI)**	**Fisher's p-value**	**Pearson χ2**	**P value**
***UNG*, chr12**												
rs80001089	T	0.80	1.00	-	0.1071	T/T	0.60	1.00			4.560	0.0327
	G	0.20	0.00			T/G	0.40	0.00	-	0.0867		
						G/G	0.00	0.00	-	1.0000		
						T/G+G/G	0.00	0.00	-	0.0867		
rs2569987	T	0.80	1.00	-	0.1071	T/T	0.60	1.00			4.560	0.0327
	C	0.20	0.00			T/C	0.40	0.00	-	0.0867		
						C/C	0.00	0.00	-	1.0000		
						T/C+C/C	0.00	0.00	-	0.0867		
rs2268406	T	0.80	0.89	2.00 (0.32-12.51)	0.6630	T/T	0.60	0.78			0.693	0.4052
	G	0.20	0.11			T/G	0.40	0.22	2.33 (0.31-17.55)	0.6285		
						G/G	0.00	0.00	-	1.0000		
						T/G+G/G	0.40	0.22	2.33 (0.31-17.55)	0.6285		
***NEIL1*, chr15**												
rs7182283	G	0.40	0.67	3.00 (0.80-11.31)	0.1192	G/G	0.20	0.56			2.574	0.1087
	T	0.60	0.33			G/T	0.40	0.22	5.00 (0.47-52.96)	0.2861		
						T/T	0.40	0.22	5.00 (0.47-52.96)	0.2861		
						G/T+T/T	0.80	0.44	5.00 (0.66-38.15)	0.1698		
***POLβ*, chr8**												
rs1012381950	**T**	**0.55**	**0.83**	**4.09 (0.89-18.72)**	**0.0860**	**T/T**	**0.10**	**0.67**			**6.537**	**0.0106**
	C	**0.45**	**0.17**			**T/C**	**0.90**	**0.33**	**18.00 (1.50-216.63)**	**0.0198**		
						C/C	**0.00**	**0.00**	**-**	**1.000**		
						T/C+C/C	**0.90**	**0.33**	**18.00 (1.50-216.63)**	**0.0198**		

***APOE*, chr19**												
rs429358	T	0.65	0.78	1.88 (0.45-7.97)	0.4848	T/T	0.40	0.56			0.460	0.4977
	C	0.35	0.22			T/C	0.50	0.44	1.56 (0.24-10.03)	1.0000		
						C/C	0.10	0.00	-	1.0000		
						T/C+C/C	0.60	0.44	1.88 (0.30-11.63)	0.6563		
hpC vs Control, TC												
	Allele frequency					**Genotype frequency**						
	Allele	**hpC**	**Control**	**OR (95% CI)**	**Fisher's p-value**	**Genotype**	**LOAD**	**Control**	**OR (95% CI)**	**Fisher's p-value**	**Pearson $\chi 2$**	**P value**
UNG, chr12												
rs2569987	T	0.88	1.00	-	0.2139	T/T	0.75	1.00			2.550	0.1103
	C	0.13	0.00			T/C	0.25	0.00	-	0.2059		
						C/C	0.00	0.00	-	1.0000		
						T/C+C/C	0.25	0.00	-	0.2059		
***POLβ*, chr8**												
rs1012381950	T	0.63	0.83	3.00 (0.61-14.86)	0.2497	T/T	0.25	0.67			2.951	0.0858
	C	0.38	0.17			T/C	0.75	0.33	6.00 (0.72-49.84)	0.1534		
						C/C	0.00	0.00	-	1.0000		
						T/C+C/C	0.75	0.33	6.00 (0.72-49.84)	0.1534		
***APOE*, chr19**												
rs429358	T	0.94	0.83	0.33 (0.03-3.58)	0.6041	T/T	0.88	0.67			1.022	0.3121
	C	0.06	0.17			T/C	0.12	0.33	0.29 (0.02-3.52)	0.5765		
						C/C	0.00	0.00	-	1.0000		
						T/C+C/C	0.12	0.33	0.29 (0.02-3.52)	0.5765		
LOAD vs hpC+Control, TC												
	Allele frequency					**Genotype frequency**						
	Allele	**LOAD**	**Control**	**OR (95% CI)**	**Fisher's p-value**	**Genotype**	**LOAD**	**Control**	**OR (95% CI)**	**Fisher's p-value**	**Pearson $\chi 2$**	**P value**
***UNG*, chr12**												
rs80001089	**T**	**0.80**	**1.00**	**-**	**0.0153**	**T/T**	**0.60**	**1.00**			**7.983**	**0.0047**

	G	**0.20**	**0.00**			**T/G**	**0.40**	**0.00**	**-**	**0.0120**		
						G/G	**0.00**	**0.00**	**-**	**1.0000**		
						T/G+G/G	**0.40**	**0.00**	**-**	**0.0120**		
rs2569987	T	0.80	0.94	4.00 (0.66-24.21)	0.1792	T/T	0.60	0.88			2.904	0.0883
	C	0.20	0.06			T/C	0.40	0.12	5.00 (0.72-34.92)	0.1535		
						C/C	0.00	0.00	-	1.0000		
						T/C+C/C	0.40	0.12	5.00 (0.72-34.92)	0.1535		
rs1610925	-	0.70	0.85	2.49 (0.65-9.56)	0.2937	-/-	0.40	0.71			2.440	0.1183
	TA	0.30	0.15			-/TA	0.60	0.29	3.60 (0.70-18.56)	0.2238		
						TA/TA	0.00	0.00	-	1.0000		
						-/TA +TA/TA	0.60	0.29	3.60 (0.70-18.56)	0.2238		
rs2268406	T	0.80	0.88	1.88 (0.41-8.51)	0.4495	T/T	0.60	0.76			0.819	0.3654
	G	0.20	0.12			T/G	0.40	0.24	2.17 (0.40-11.74)	0.4147		
						G/G	0.00	0.00	-	1.0000		
						T/G+G/G	0.40	0.24	2.17 (0.40-11.74)	0.4147		
***NEIL1*, chr15**												
rs7182283	G	0.40	0.65	2.75 (0.88-8.58)	0.0955	G/G	0.20	0.47			1.977	0.1597
	T	0.60	0.35			G/T	0.40	0.35	2.67 (0.36-19.71)	0.6285		
						T/T	0.40	0.18	5.33 (0.62-45.99)	0.1618		
						G/T+T/T	0.80	0.53	3.56 (0.58-21.92)	0.2305		
***POLβ*, chr8**												
rs1012381950	T	0.55	0.74	2.27 (0.71-7.28)	0.2330	T/T	0.10	0.50			3.891	0.0485
	C	0.45	0.26			T/C	0.90	0.50	8.00 (0.82-77.82)	0.0912		
						C/C	0.00	0.00	-	1.0000		
						T/C+C/C	0.90	0.50	8.00 (0.82-77.82)	0.0912		
***APOE*, chr19**												
rs429358	T	0.65	0.79	2.08 (0.60-7.17)	0.3368	T/T	0.40	0.59			0.894	0.3445
	C	0.35	0.21			T/C	0.50	0.41	1.79 (0.35-9.13)	0.6828		
						C/C	0.10	0.00	-	0.3333		
						T/C+C/C	0.60	0.41	2.14 (0.44-10.53)	0.4401		

4.6.1. *APOE* allele and genotype frequencies and comparison with BER genes in LOAD-CE and LOAD-TC samples

APOE allele and genotype frequencies of post-mortem brain tissues are presented in Table 4.13. The *APOE* ε4 allele frequency was 0.35 for both LOAD-CE and TC and 0.22 for both control-CE and TC but it did not show statistically significant relation with LOAD (p=0.4848). The *APOE* ε4 allele frequency of hpC individuals was lower according to control group; 0.13 and 0.19 in CE and TC respectively. The *APOE* ε2 allele was not observed in any of the brain tissue sample groups. Also ε3/ε4 genotype frequency was same for both LOAD (0.50) and control group (0.44) in both CE and TC samples and it did not show statistically significant relation with LOAD. The *APOE* ε4/ε4 LOAD genotype frequency was 0.10 with no statistically significant relation in both TC and CE and this genotype was not found in hpC group.

Table 4.13. Table 4.6 Allele and genotype frequencies of *APOE* ε2, ε3 and ε4 in LOAD-CE and TC s**amples.** Bold; p-value <0

rs429358-rs7412					
LOAD vs Control, CE					
		LOAD	Control	OR (95% CI)	Fisher's p-value
Allele					
ε2	T-T	0.00	0.00	-	1.0000
ε3	T-C	0.65	0.78	(REF)	(REF)
ε4	C-C	0.35	0.22	1.88 (0.45-7.97)	0.4848
Genotype					
ε3/ε4	TT-TT	0.50	0.44	1.56 (0.24-10.03)	1.0000
ε4/ε4	TT-TC	0.10	0.00	-	1.0000
ε2/ε2	TC-TC	0.00	0.00	-	1.0000
ε2/ε3	TT-CC	0.00	0.00	-	1.0000
ε2/ε4	TC-CC	0.00	0.00	-	1.0000
ε3/ε3	CC-CC	0.40	0.56	(REF)	(REF)
hpC vs Control, CE					
		hpC	Control	OR (95% CI)	Fisher's p-value
Allele					
ε2	T-T	0.00	0.00	-	1.0000
ε3	T-C	0.88	0.78	(REF)	(REF)
ε4	C-C	0.13	0.22	0.50 (0.08-3.19)	0.6602

Genotype					
ε3/ε4	TT-TT	0.25	0.44	0.42 (0.05-3.31)	0.6199
ε4/ε4	TT-TC	0.00	0.00	-	1.0000
ε2/ε2	TC-TC	0.00	0.00	-	1.0000
ε2/ε3	TT-CC	0.00	0.00	-	1.0000
ε2/ε4	TC-CC	0.00	0.00	-	1.0000
ε3/ε3	CC-CC	0.75	0.56	(REF)	(REF)
LOAD vs hpC+Control, CE					
		LOAD	Control	OR (95% CI)	Fisher's p-value
Allele					
ε2	T-T	0.00	0.00		
ε3	T-C	0.65	0.82	(REF)	(REF)
ε4	C-C	0.35	0.18	2.51 (0.70-8.98)	0.1936
Genotype					
ε3/ε4	TT-TT	0.50	0.35	2.29 (0.44-11.92)	0.4185
ε4/ε4	TT-TC	0.10	0.00	-	0.3125
ε2/ε2	TC-TC	0.00	0.00	-	1.0000
ε2/ε3	TT-CC	0.00	0.00	-	1.0000
ε2/ε4	TC-CC	0.00	0.00	-	1.0000
ε3/ε3	CC-CC	0.40	0.65	(REF)	(REF)
LOAD vs Control, TC					
		LOAD	Control	OR (95% CI)	Fisher's p-value
Allele					
ε2	T-T	0.00	0.00	-	1.0000
ε3	T-C	0.65	0.78	(REF)	(REF)
ε4	C-C	0.35	0.22	1.88 (0.45-7.97)	0.4848
Genotype					
ε3/ε4	TT-TT	0.50	0.44	1.56 (0.24-10.03)	1.0000
ε4/ε4	TT-TC	0.10	0.00	-	1.0000
ε2/ε2	TC-TC	0.00	0.00	-	1.0000
ε2/ε3	TT-CC	0.00	0.00	-	1.0000
ε2/ε4	TC-CC	0.00	0.00	-	1.0000
ε3/ε3	CC-CC	0.40	0.56	(REF)	(REF)
hpC vs Control, TC					
		hpC	Control	OR (95% CI)	Fisher's p-value
Allele					
ε2	T-T	0.00	0.00	-	1.0000
ε3	T-C	0.81	0.78	(REF)	(REF)
ε4	C-C	0.19	0.22	0.81 (0.15-4.32)	1.0000
Genotype					
ε3/ε4	TT-TT	0.38	0.44	0.75 (0.11-5.24)	1.0000
ε4/ε4	TT-TC	0.00	0.00	-	1.0000
ε2/ε2	TC-TC	0.00	0.00	-	1.0000
ε2/ε3	TT-CC	0.00	0.00	-	1.0000
ε2/ε4	TC-CC	0.00	0.00	-	1.0000
ε3/ε3	CC-CC	0.62	0.56	(REF)	(REF)
LOAD vs hpC+Control, TC					

		LOAD	Control	OR (95% CI)	Fisher's p-value
Allele					
ε2	T-T	0.00	0.00	-	1.0000
ε3	T-C	0.65	0.79	(REF)	(REF)
ε4	C-C	0.35	0.21	1.12 (0.46-2.75)	0.3368
Genotype					
ε3/ε4	TT-TT	0.50	0.42	1.79 (0.35-9.13)	0.6828
ε4/ε4	TT-TC	0.10	0.00	-	0.3333
ε2/ε2	TC-TC	0.00	0.00	-	1.0000
ε2/ε3	TT-CC	0.00	0.00	-	1.0000
ε2/ε4	TC-CC	0.00	0.00	-	1.0000
ε3/ε3	CC-CC	0.40	0.58	(REF)	(REF)

The relation between statistically significantly related *UNG* rs2569987 for LOAD-CE, *APOE* rs405509 for hpC-CE and *UNG* rs80001089 for LOAD-TC and APOE ε4 carrying status, was investigated (Table 4.14). Since no ε2/ε4 genotype carrier was found, none of the individuals was excluded from the calculations. *APOE* ε4 carrying status did not have a statistically significant effect on LOAD risk for both *UNG* rs2569987 (LOAD-CE) and *UNG* rs80001089 (LOAD-TC) carriers or non-carriers. And also *UNG* rs2569987 (LOAD-CE) and *UNG* rs80001089 (LOAD-TC) carrying status did not have a statistically significant effect on LOAD for *APOE* ε4 carriers or non-carriers. For *APOE* ε4 non-carriers *APOE* rs405509 had a statistically significant association in hpC group (p=0.0476).

Table 4.14. Statistically significantly related variants and *APOE* ε4 carrying status interaction in LOAD-CE and TC samples. Bold, p-value <0; +, carrier; -, non-carrier

LOAD vs Control, CE					
		LOAD	**Control**	**OR (95% CI)**	**Fisher's p-value**
APOE ε4 (+)	rs2569987 (+)	0.20	0.00	-	0.1667
	rs2569987 (-)	0.20	0.56		
APOE ε4 (-)	rs2569987 (+)	0.20	0.00	-	0.4667
	rs2569987 (-)	0.40	0.44		
rs2569987 (+)	APOE ε4 (+)	0.20	0.00	-	-
	APOE ε4 (-)	0.20	0.00		
rs2569987 (-)	APOE ε4 (+)	0.20	0.56	0.40 (0.05-3.42)	0.6084
	APOE ε4 (-)	0.40	0.44		
hpC vs Control, CE					
		hpC	**Control**	**OR (95% CI)**	**Fisher's p-value**
APOE ε4 (+)	rs405509 (+)	0.125	0.00	-	0.3750
	rs405509 (-)	0.250	0.56		
APOE ε4 (-)	**rs405509 (+)**	**0.500**	**0.00**	**-**	**0.0476**
	rs405509 (-)	**0.125**	**0.44**		
rs405509 (+)	APOE ε4 (+)	0.125	0.00	-	-

rs405509 (-)	APOE ε4 (-)	0.500	0.00	1.60 (0.10-24.70)	1.0000
	APOE ε4 (+)	0.250	0.56		
	APOE ε4 (-)	0.125	0.44		
LOAD vs hpC+Control, CE					
		LOAD	**Control**	**OR (95% CI)**	**Fisher's p-value**
APOE ε4 (+)	rs2569987 (+)	0.20	0.06	11.00 (0.65-187.18)	0.1357
	rs2569987 (-)	0.20	0.65		
APOE ε4 (-)	rs2569987 (+)	0.20	0.00	-	0.4545
	rs2569987 (-)	0.40	0.29		
rs2569987 (+)	APOE ε4 (+)	0.20	0.06	-	1.0000
	APOE ε4 (-)	0.20	0.00		
rs2569987 (-)	APOE ε4 (+)	0.20	0.65	0.23 (0.03-1.68)	0.1778
	APOE ε4 (-)	0.40	0.29		
LOAD vs hpC+Control, TC					
		LOAD	**Control**	**OR (95% CI)**	**Fisher's p-value**
APOE ε4 (+)	rs80001089 (+)	0.10	0.00	-	0.1667
	rs80001089 (-)	0.10	0.59		
APOE ε4 (-)	rs80001089 (+)	0.30	0.00	-	0.2000
	rs80001089 (-)	0.50	0.41		
rs80001089 (+)	APOE ε4 (+)	0.10	0.00	-	-
	APOE ε4 (-)	0.30	0.00		
rs80001089 (-)	APOE ε4 (+)	0.10	0.59	0.14 (0.01-1.47)	0.1550
	APOE ε4 (-)	0.50	0.41		

4.6.2. Genetic models of statistically significant LOAD-CE and TC variants

In Table 4.15 genetic model comparison of LOAD-CE and TC variants were presented. Both *UNG* rs80001089 when compared with control+hpC and *POLβ* rs1012381950 in LOAD-TC when compared with only control were fitted better to dominant model. But for LOAD-CE *UNG* rs2569987 when compared with only control was fitted better to additive model.

Table 4.15. Genetic models of statistically significantly related *UNG* variants in LOAD-TC and CE samples. Bold, p-value <0

		Additive Model		**Dominant Model**		**Recessive Model**	
	Tissue type	**OR (95% CI)**	**Fisher's p-value**	**OR (95% CI)**	**Fisher's p-value**	**OR (95% CI)**	**Fisher's p-value**
***UNG* rs80001089**	TC	-	0.0153	-	0.0120	-	1.0000
***POLβ* rs1012381950**	TC	-	0.0860	-	0.0198	-	1.0000
***UNG* rs2569987**	CE	-	0.0216	-	0.0351	-	1.0000

4.6.3. Epistatic relations of statistically significant LOAD-TC samples

In LOAD-TC samples the epistatic relation between *UNG* rs80001089 and *NEIL1* rs7182283 was found statistically significantly related to LOAD (p=0.0410) (Table 4.16).

Table 4.16. Epistatic relation of LOAD-TC significant variants. Frequencies in LOAD and Control group with their OR and p values were showed. Bold, p-value <0

	Variations	LOAD Frequency	Control Frequency	OR (95% CI)	Fisher`s p-value
UNG-NEIL1	**rs80001089-rs7182283**	**0.300**	**0.000**	**-**	**0.0410**

4.7. Comparison of Statistically Significant Variations Between Study Groups

The comparison of statistically significantly related variants is presented in Table 4.17. In LOAD-blood group, *UNG* gene variants rs1610925 TG insertion, rs2268406 G allele, rs80001089 G allele, rs1018782 A allele and rs1018783 A allele and *APOE* gene rs769449 A allele and rs429358 C allele were statistically significantly related. *UNG* gene rs80001089 G allele was also statistically significantly related in LOAD-TC group (Table 4.17 and Figure 4.7A). *POLβ* rs1012381950 T/C genotype was statistically significantly related in LOAD-TC group. In LOAD-CE group *UNG* rs2569987 C allele was statistically significantly related with LOAD when compared with controls and control+hpCs but this variant did not show statistical significance in blood or TC group (Table 4.17 and Figure 4.7A). *APOE* rs405509 G allele was statistically significantly related with hpC group in CE samples. Except *APOE* gene variants rs429358 and rs405509, all statistically significantly related variants are located in intron regions of relevant genes. rs429358 is located in 4th exon of *APOE* gene while rs405509 is in upstream region (Figure 4.7B).

Table 4.17. Statistically significantly related variations in LOAD, control and hpC groups of blood, TC and CE samples.

			Frequency			Fisher`s p-value			
Gene	**Variant**	**Allele**	**LOAD**	**Control**	**hpC**	**LOAD vs Control**	**hpC vs Control**	**LOAD vs hpC + Control**	**Sample**
			0.197	0.087	-	**0.0005**	-	-	Blood
UNG	rs1610925	TG	0.300	0.222	0.063	0.7190	0.3402	0.2937	TC
			0.182	0.200	0.045	1.0000	0.1745	0.5297	CE
			0.164	0.082		**0.0052**	-	-	Blood
UNG	rs2268406	G	0.200	0.111	0.125	0.6630	1.0000	0.4495	TC
			0.182	0.050	0.045	0.6653	1.0000	0.2204	CE
			0.109	0.046	-	**0.0129**	-	-	Blood
UNG	rs80001089	G	0.200	0.000	0.000	0.1071	NA	**0.0153**	TC
			0.091	0.000	0.000	0.3465	NA	0.0437	CE
			0.152	0.082	-	**0.0184**	-	-	Blood
UNG	rs1018782	A	0.200	0.167	0.187	1.0000	1.0000	1.0000	TC
			0.000	0.000	0.000	NA	NA	NA	CE
			0.134	0.077	-	**0.0406**	-	-	Blood
UNG	rs1018783	A	0.200	0.167	0.125	1.0000	1.0000	1.0000	TC
			0.000	0.000	0.000	NA	NA	NA	CE
			0.078	0.061	-	0.5046	-	-	Blood
UNG	rs2569987	C	0.200	0.000	0.125	0.1071	0.2139	0.1792	TC
			0.273	0.000	0.200	**0.0216**	0.4890	**0.0162**	CE
			0.134	0.026	-	**0.0001**	-	-	Blood
APOE	rs769449	A	0.100	0.111	0.125	1.0000	1.0000	1.0000	TC
			0.091	0.136	0.045	0.6560	0.3327	1.0000	CE
			0.189	0.061	-	**0.0001**	-	-	Blood
APOE	rs429358	C	0.350	0.222	0.188	0.4848	0.6041	0.3368	TC
			0.318	0.200	0.136	0.4913	0.6909	0.3522	CE
			0.472	0.444	-	0.5404	-	-	Blood
APOE	rs405509	G	0.450	0.500	0.250	1.0000	0.1717	0.7753	TC
			0.045	0.000	0.273	1.0000	**0.0216**	0.4062	CE
Gene	**Variant**	**Genotype**	**LOAD**	**Control**	**hpC**	**LOAD vs Control**	**hpC vs Control**	**LOAD vs hpC +Control**	**Sample**
			0.005	0.00	-	1.0000	-	-	Blood
POLβ	rs1012381950	T/C	0.900	0.333	0.750	**0.0198**	0.1534	0.0912	TC
			0.091	0.010	0.091	1.0000	1.0000	1.0000	CE

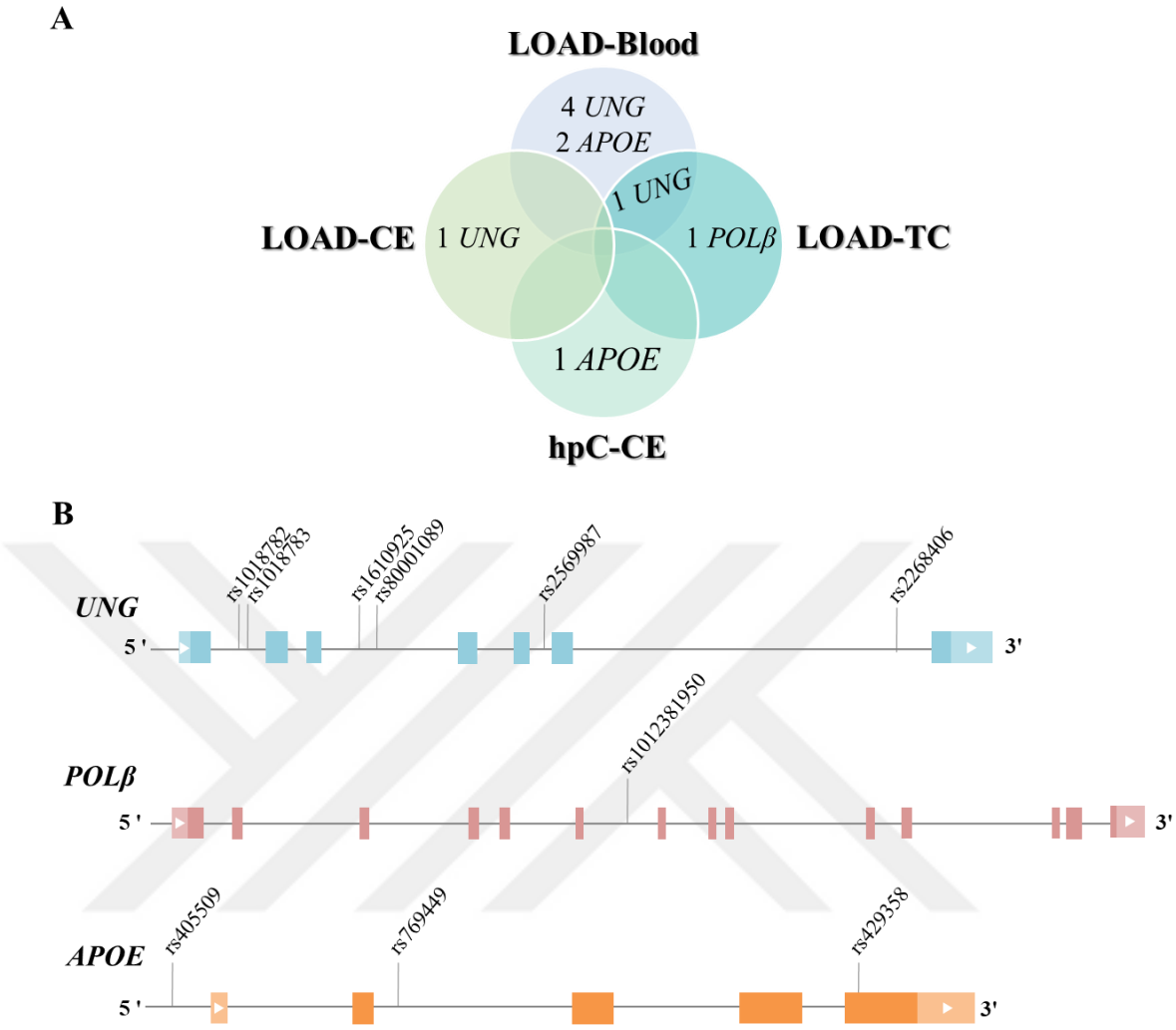

Figure 4.7. Statistically significantly related variants comparison in all sample groups. A) Specific and shared variants between LOAD-blood, CE, TC and hpC samples are shown. B) Statistically significantly related variants locations on gene structure are shown.

4.8. Gene Expression Analysis of *UNG* in LOAD-blood Samples

The effects of statistically significantly related *UNG* variants on gene expression levels were analyzed using qRT-PCR. Six LOAD patients and 4 control individuals were used for the analysis. Four of the 6 LOAD-blood patients carry 3 (rs1610925, rs2268406, and rs80001089) of the significantly related *UNG* gene variants and 2 of them carry 4 (rs1610925, rs2268406, rs1018782 and rs101878) significantly related *UNG* gene variants. Control individuals do not carry the significantly related *UNG* gene variants. The mean $2^{-\Delta\Delta CT}$ value of the LOAD-blood patients and control individuals

were as follows; 0.951±0.394 and 1.093±0.697. Although the gene expression level was higher in control group, this difference was not significant (p=0.3524) (Figure 4.8).

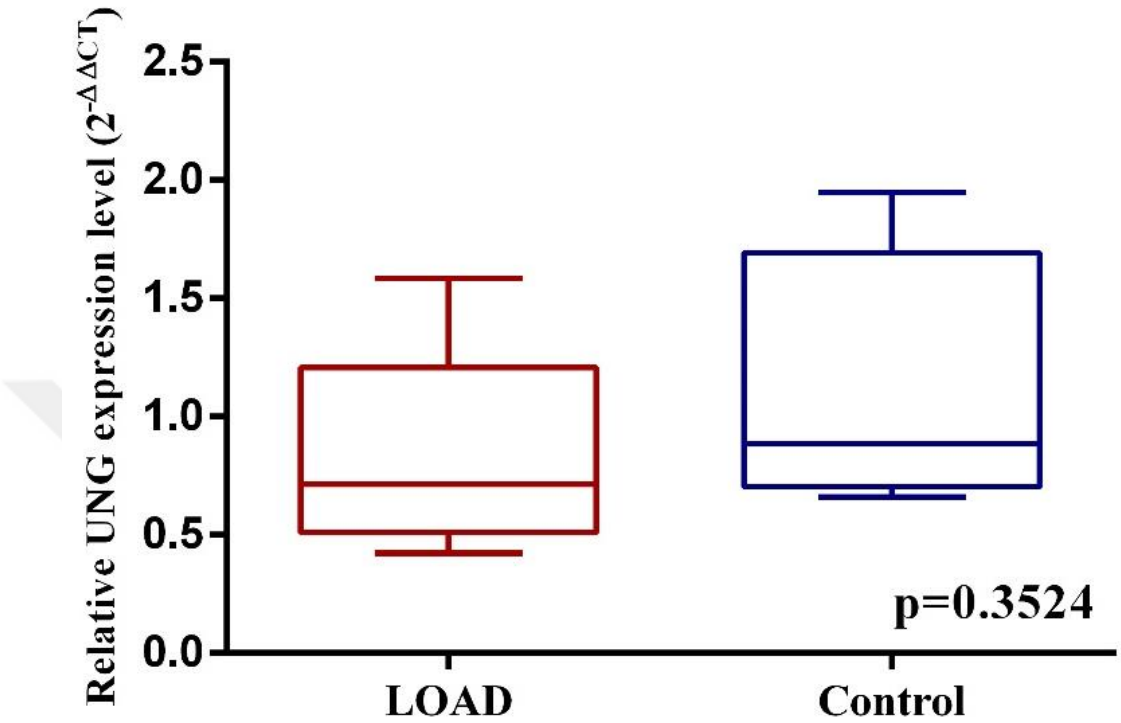

Figure 4.8. UNG gene expression levels of LOAD patients and control individuals.

4.9. Common Variations in Study Population

Common variations in study population was summarized in Table 4.18 (maf>0.05). The variants were listed in order from the most common to the least. The MAF values were compared with; all 1000G phase 3 individuals (ALL); African (AFR), American (AMR), East Asian (EAS), European (EUR) and South Asian (SAS). In general, the MAFs of the variants in Turkish population showed similarity to other populations except AFR population. *UNG* variant T/C located on 109,541,016 coordinate was a novel variant, which was not reported before. It was found in 8% of the Turkish population.

Table 4.18. Common variations in Turkish population

Gene	Variation	MA	Genomic location	MAF in Turkish population	ALL	AFR	AMR	EAS	EUR	SAS
NEIL1	rs7182283	G	intronic	0.47	0.53	0.74	0.37	0.48	0.52	0.44
APOE	rs405509	G	upstream	0.44	0.53	0.76	0.5	0.33	0.52	0.45
UNG	rs2541886	T	intronic	0.44	0.25	0.04	0.29	0.19	0.50	0.30
NEIL1	rs5745920	T	downstream	0.39	0.39	0.58	0.24	0.34	0.28	0.39
UNG	rs246079	G	intronic	0.36	0.62	0.75	0.66	0.71	0.41	0.53
		A		0.64	0.38	0.25	0.34	0.29	0.59	0.47
UNG	rs3219211	C	intronic	0.18	0.23	0.20	0.15	0.30	0.17	0.30
UNG	rs3219243	C	intronic	0.18	0.23	0.22	0.15	0.3	0.17	0.27
APOE	rs449647	T	upstream	0.18	0.20	0.33	0.26	0.30	0.17	0.19
POLβ	rs35172933	AGG	intronic	0.17	0.35	0.84	0.14	0.17	0.12	0.28
POLβ	rs2976239	C	intronic	0.17	0.13	0.39	0.05	0.01	0.04	0.02
POLβ	rs2976238	T	intronic	0.16	0.30	0.72	0.10	0.17	0.08	0.25
POLβ	rs2979896	C	intronic	0.15	0.31	0.74	0.10	0.17	0.09	0.25
APOE	rs429358	C	exonic	0.15	0.15	0.27	0.10	0.09	0.16	0.09
POLβ	rs6474390	T	intronic	0.15	0.31	0.74	0.10	0.17	0.09	0.25
UNG	rs2268406	G	intronic	0.14	0.14	0.22	0.12	0.04	0.16	0.15
POLβ	rs2979895	G	intronic	0.14	0.30	0.72	0.10	0.17	0.09	0.25
POLβ	rs2953983	G	intronic	0.14	0.30	0.72	0.10	0.17	0.09	0.25
POLβ	rs3136717	C	intronic	0.14	0.34	0.77	0.14	0.17	0.12	0.28
POLβ	rs3136718	A	intronic	0.13	0.31	0.74	0.10	0.17	0.09	0.25
UNG	rs1018782	G	intronic	0.13	0.21	0.44	0.15	0.04	0.16	0.15
POLβ	rs2953993	T	intronic	0.13	0.30	0.72	0.10	0.17	0.09	0.25
UNG	rs3219235	T	intronic	0.13	0.23	0.20	0.15	0.30	0.17	0.30
UNG	rs1610925	TA	intronic	0.12	0.25	0.42	0.18	0.10	0.22	0.23
POLβ	rs3136790	C	intronic	0.12	0.34	0.79	0.14	0.17	0.12	0.28
POLβ	rs2272615	G	intronic	0.12	0.33	0.77	0.13	0.17	0.12	0.28
POLβ	rs3136781	C	intronic	0.12	0.34	0.79	0.14	0.17	0.13	0.28
POLβ	rs2976244	T	intronic	0.12	0.31	0.74	0.10	0.17	0.09	0.25
UNG	rs1018783	A	intronic	0.12	0.19	0.40	0.15	0.04	0.16	0.15
POLβ	rs3136794	G	intronic	0.11	0.33	0.76	0.14	0.17	0.12	0.28
NEIL1	rs11634109	C	intronic	0.10	0.10	0.07	0.11	0.03	0.24	0.05
APOE	rs769449	A	intronic	0.10	0.06	0.01	0.08	0.08	0.12	0.06
POLβ	rs3136738	T	intronic	0.10	0.31	0.74	0.10	0.17	0.09	0.25
POLβ	rs2953994	G	intronic	0.10	0.31	0.74	0.10	0.17	0.09	0.25
UNG	rs80001089	G	intronic	0.09	0.03	0.00	0.05	0.00	0.09	0.03
POLβ	rs3136811	G	intronic	0.09	0.31	0.74	0.10	0.17	0.09	0.25
POLβ	rs3136722	T	intronic	0.09	0.13	0.40	0.05	0.01	0.04	0.02
UNG	109,541,016	C	intronic	0.08	-	-	-	-	-	-
NEIL1	rs5745925	-	exonic	0.08	0.03	0.01	0.05	0.00	0.08	0.02
POLβ	rs35609234	-	intronic	0.08	0.34	0.77	0.14	0.17	0.12	0.28
POLβ	rs11990332	G	intronic	0.08	0.34	0.77	0.14	0.17	0.12	0.28
POLβ	rs59423074	C	intronic	0.08	0.31	0.74	0.10	0.17	0.09	0.25
POLβ	rs3136780	A	intronic	0.08	0.30	0.72	0.10	0.17	0.09	0.25
POLβ	rs3136793	G	intronic	0.08	0.30	0.72	0.10	0.17	0.09	0.25
POLβ	rs2976240	G	intronic	0.08	0.06	0.15	0.04	0.01	0.04	0.02
UNG	rs2569987	C	intronic	0.07	0.06	0.03	0.10	0.00	0.16	0.04
POLβ	rs3136806	G	intronic	0.07	0.22	0.42	0.17	0.14	0.10	0.22
POLβ	rs141279009	A	intronic	0.07	0.01	0.00	0.01	0.00	0.01	0.03
UNG	rs246078	A	intronic	0.06	0.00	0.00	0.00	0.00	0.02	0.00

Gene	Variation	MA	Genomic location	MAF in Turkish population	ALL	AFR	AMR	EAS	EUR	SAS
APOE	rs769446	C	upstream	0.06	0.06	0.03	0.04	0.11	0.08	0.07
NEIL1	rs5745916	A	intronic	0.05	0.01	0.00	0.01	0.00	0.03	0.01
POLβ	rs3136748	T	intronic	0.05	0.18	0.33	0.05	0.16	0.04	0.23
POLβ	rs2307158	T	upstream	0.05	0.22	0.48	0.05	0.16	0.05	0.23

5. DISCUSSION AND CONCLUSION

LOAD is the most frequent type of dementia and it has the highest prevalence in the elderly. The distribution of histopathological signatures, APs and NFTs are well understood but the mechanism underlying the accumulation of these signatures remains elusive. The environmental and genetic risk factors causing the difference of susceptibility level to LOAD and molecular mechanism leading to neuronal loss also remain elusive. There are no determined specific treatment models, biomarkers and risk factors of LOAD (2,29). The effects of oxidative DNA damage repair (BER pathway) deficiency on LOAD progression and etiology have been demonstrated (5-8). But these studies couldn't reveal a full explanation to, in LOAD patients, whether the BER gene variants are the underlying reason for decreased activity and/or protein levels of BER or they are relevant to the risk and progression of the disease. This thesis study demonstrates for the first time that five *UNG* gene non-coding variants; rs1610925, rs2268406, rs80001089, rs1018782 and rs1018783 are statistically significantly related to LOAD in Turkish blood samples ($p<0.05$). *UNG* rs80001089 is also statistically significantly related to LOAD in Brazilian post-mortem TC samples ($p<0.05$). *POLβ* rs1012381950 and *UNG* rs256998 gene variants were found statistically significantly related to LOAD only in Brazilian post-mortem CE and TC samples, respectively ($p<0.05$). CE and TC regions of the same Brazilian subjects did not have a common statistically significant variant.

Decreased level of UNG protein and enzyme activity was shown in several studies using postmortem brain tissues of LOAD patients (5, 17, 50). The location of *UNG* gene variants rs1018782 (CCAT box's 4 bp downstream) and rs1018783 (in Yi element) are in *UNG1* promoter B (97). In normal and cancer cell lines, mapping *UNG* gene variants revealed, rs1018782 (position 1034) and rs1018783 (position 1082) are genetically linked. But even though they are located in the promoter region, they did not affect transcriptional activity (106). In this thesis study, both variants were also appeared together in blood samples (~88%) and post-mortem TC samples (33%) but did not appear in post-mortem CE samples at all. These gene variants showed

statistically significant differences ($p<0.05$) in allele and genotype frequencies of LOAD-blood samples, but not in CE and TC samples. The epistatic relation (OR = 1.82, p=0.723) did not increase their individual statistical significance (OR = 2.01, p=0.0184; OR = 1.86, p=0.0406, respectively), therefore did not contribute to LOAD risk factor, as expected. The effects of non-coding *UNG* rs1610925, rs2268406, rs80001089 and rs2569987 variants on activity or expression level of UNG have not been determined. Rs80001089 and rs2569987 variants were statistically significantly related to LOAD in TC and CE samples, respectively. The UNG activity of the postmortem brain tissues, used for this targeted NGS study, was determined. The UNG2 and UNG1 activities were decreased TC samples but only UNG2 activity was decreased in CE samples. The UNG protein level did not change in both tissues, suggesting that UNG protein phosphorylation might be the underlying reason for decreased UNG activity (50). In agreement with these results, in this thesis study, no statistically significant coding gene variant, which can affect the UNG protein level was detected. This study results suggest that LOAD brain tissue's statistically significant *UNG* variants may not affect the protein level. Although there is no *UNG* gene variant identified, that affects protein function, this study is the first study for revealing the association of *UNG* gene variants with protein level and function in LOAD postmortem brain tissues, using targeted NGS (covering the whole gene). Thus *UNG* or other BER gene variants showed no statistically significant relation with hpC individuals, whose UNG protein levels did not decrease (50). The knowledge on the relation between *UNG* gene variants and human diseases is very little. In Taiwan's Han Chinese population, the *UNG* rs246079 A/G variant was related to susceptibility to rheumatoid arthritis and increased risk of lung cancer (107,108). Contrarily in a Chinese population, G/A variant was related to a decreased risk of esophageal cancer (109). In this thesis study, the *UNG* rs246079 variant was found as a common variant in the Turkish study population with a 0.36 MAF (Table 4.18).

Studies using NEIL1 and POLβ knockout mice showed that this depletion led to AD pathogenesis and phenotype development. Also, studies with LOAD patients showed a decrease in NEIL1 and POLβ protein levels and activity. These results indicate the importance of NEIL1 and POLβ for AD (9-12). In this thesis

study, *NEIL1* or *POLβ* variants did not statistically significantly related to LOAD in Turkish population blood samples, which indicate other factors, such as translational or post-transcriptional modifications, can be the reason of decreased protein level and activities in LOAD. In LOAD patients and healthy controls, *NEIL1* promoter methylation status did not affect the downregulation of NEIL1 (69). In a SNP genotyping study, gene-gene interactions of *OGG1* rs1052133 and *NEIL1* rs4462560 was found to be statistically significantly related with LOAD (OR = 2.24, 95% CI = 1.36–3.91, p = 0.041) (73). The primers used for this thesis study do not cover the location of *NEIL1* rs4462560 (chr15:75,647,964) variant (Table 3.1, Figure 3.1). Unlike the Turkish population, in Brazilian post-mortem TC samples of LOAD, *UNG* rs80001089 (p = 0.0153) and *NEIL1* rs7182283 (p = 0.0955) epistatic relation was found statistically significant with LOAD (p = 0.0410) yet alone *NEIL1* rs7182283 did not show statistical significance. This result points to the relationship between LOAD and the combinatory effect of *UNG* and *NEIL1*. In blood samples, *NEIL1* rs7182283 is a common variant (Table 4.18) and its epistasis with other variants did not contribute to LOAD risk. In LOAD CE samples, the *POLβ* rs1012381950 T/C genotype was statistically significantly related to LOAD (p = 0.0198), suggesting its association with LOAD.

So far only accepted genetic risk factor for LOAD is *APOE* ε4 allele and ε3 is the most frequently found allele. Also in this thesis study, LOAD and *APOE* ε4 association and frequency of ε3 were confirmed. In the Turkish population, the frequency of *APOE* ε4 in LOAD was studied by two studies so far (110,111). In the Turkish population, to associate the *APOE* ε4 with LOAD risk, this is the only study that sequenced the whole gene using NGS, as far as we know. In other Turkish population studies, *APOE* ε4 allele frequency was 11.4% and 17.2% which is smaller than our study population (18.94%) (110,111). Age, sex, geographical location and also sample size may be the reason for the difference between frequencies. The Turkish population *APOE* ε4 allele frequency is lower than other studied populations such as Caucasian (36.7%), African-American (32.3%), Hispanic (19.2%) and Japanese (27.8%) (www.AlzGene.org). Carrying *APOE* ε4/ε4 genotype increases LOAD susceptibility to 10 times, while ε3/ε4 genotype increases 3 times (39). *APOE* ε3/ε4

genotype frequency was higher according to *APOE* ε4/ε4 in this study indicating that in Turkish population, LOAD main risk factor is *APOE* ε3/ε4 genotype. In the Turkish population, *APOE* ε3/ε4 genotype contributed to the LOAD risk stronger (OR =4.07) than other populations, such as Hispanic (OR = 2.2), Caucasian (OR = 2.7), African-American (OR = 1.1) but weaker than Japanese (OR = 5.6). On the other hand, *APOE* ε4/ε4 was a weaker contributor to the LOAD risk factor when compared with other populations (39). Gene-gene interactions of *APOE* with and *PSEN1* rs17125721 and *GAB2* rs2373115 was shown to be increasing the LOAD risk (112,113). But in this thesis study, *APOE* ε4 and statistically significantly related *UNG* gene carrying status was inversely related. In *APOE* ε4 non-carriers, *UNG* rs1610925 (OR=2.58), rs2268406 (OR=2.21), rs80001089 (OR=2.53), rs1018782 (OR=2.01), rs1018783 (OR=1.86) variants risk contribution was increased (OR= 4.00, 3.83, 6.03, 3.11, 2.78, respectively). Also in these statistically significantly related *UNG* variants non-carriers *APOE* ε4 (OR=3.58) carrying increased the risk contribution (OR=6.88, 6.88, 6.30, 6.11, 6.27, respectively) (Table 4.14).

In this thesis study, the *APOE* ε3 allele was found to be the most frequent *APOE* allele in both CE and TC samples. Also, studies focusing on *APOE* and LOAD relation in the Brazilian population, have shown that *the APOE* ε3 allele is the most frequent (114-117). The comparison of *APOE* ε4 allele with age-matched control using LOAD patient's hippocampus in Brazilian revealed no statistically significant association (118,119). In addition to these studies, in this study, the frequency of *APOE* ε4 allele in LOAD patients (35%) was higher according to age-matched controls (22%) but this difference did not show statistical significance. There were also some studies that showed a higher ε4 allele frequency when compared with LOAD in post-mortem brain tissues but no statistical significance was reported (120-123). These differences in the distribution of *APOE* ε4 allele may be due to the sample size, difference in post-mortem brain tissue region and ethnic background.

In Turkish population blood samples, *APOE* rs769449 (OR=5.90) and rs429358 (OR= 3.58) variants showed statistically significant association (p=0.0001). The

epistatic interaction of these gene variants showed a higher contribution to LOAD risk (OR=6.12, p=0.0001). The rs769449 variant was showed to be in strong linkage disequilibrium with APOE ε2/ε3/ε4 (124). The statistically significant association of the rs769449 variant was suggested to be related to its relation with *APOE* ε4 (124,125). This thesis study has also supported the association between them. In *APOE* ε4 carriers rs769449 carrying status increased the risk contribution to LOAD (OR=6.09, 95% CI = 1.42–26.17, p = 0.0131). It was suggested that rs769449 may affect *APOE* gene region epigenetic state that influences protein concentration and transcription levels (124,125). In the Turkish population, *APOE* rs769449 and rs429358 were also found as common variants (MAF 0.10 and 0.15, respectively). *APOE* rs429358 was found as common variant in all populations with the highest frequency in Africans. On the other hand, *APOE* rs769449 was also found as common variant in all populations but interestingly except Africans (Table 4.18).

In the Turkish population sporadic PD samples, *APOE* upstream variant rs769446 C allele and C/T genotype were found to be statistically significantly important (p=0.0001). To the best of our knowledge, this association is showed for the first time using targeted NGS. *APOE* rs769446 gene variant did not show a statistically significant association with LOAD in Turkish or Brazilian population samples. Association of *APOE* rs769446 and LOAD was examined by several studies and in general, also no association was found with LOAD and the association situation did not change with *APOE* ε4 carrier or non-carrier situation (114, 126-130). In one of these studies, the effect of different genetic models of rs769446 variant was studied and only allelic model for C was found, statistically significantly associated with LOAD (OR = 1.27, 95 % CI = 1.114–1.449, P < 0.0001) (130). *APOE* rs769446 was found as a common variant in the Turkish population (MAF=0.06) in line with other human populations.

Taking everything into account, our results suggest that in APOE ε4 non-carriers, statistically significant BER gene variants may be associated with LOAD risk. Although these variants are statistically significant, they do not affect protein structure.

This suggests that post-transcriptional or–translational modifications may be the reason for decreased protein levels and activities of BER genes in LOAD. The relationship between BER variants and LOAD risk needs to be confirmed by an increased sample size in further studies. Also, these variants are unique to LOAD when compared with sporadic PD. But, this needs to be confirmed by studying these BER variants with other neurodegenerative disorders. For our comprehension of the BER effect on LOAD, sporadic PD and other neurodegenerative disorders progression, these results would open a new door for BER and LOAD association.

6. REFERENCES

1. Honea RA, Vidoni ED, Swerdlow RH, Burns JM, Alzheimer's Disease Neuroimaging I. Maternal family history is associated with Alzheimer's disease biomarkers. J Alzheimers Dis. 2012;31(3):659-68.

2. Braak H, Braak E. Neuropathological stageing of Alzheimer-related changes. Acta Neuropathol. 1991;82(4):239-59.

3. Poewe W, Seppi K, Tanner CM, Halliday GM, Brundin P, Volkmann J, et al. Parkinson disease. Nat Rev Dis Primers. 2017;3:17013.

4. Ascherio A, Schwarzschild MA. The epidemiology of Parkinson's disease: risk factors and prevention. Lancet Neurol. 2016;15(12):1257-72.

5. Weissman L, Jo DG, Sorensen MM, de Souza-Pinto NC, Markesbery WR, Mattson MP, et al. Defective DNA base excision repair in brain from individuals with Alzheimer's disease and amnestic mild cognitive impairment. Nucleic Acids Res. 2007;35(16):5545-55.

6. Lovell MA, Gabbita SP, Markesbery WR. Increased DNA oxidation and decreased levels of repair products in Alzheimer's disease ventricular CSF. J Neurochem. 1999;72(2):771-6.

7. Lovell MA, Soman S, Bradley MA. Oxidatively modified nucleic acids in preclinical Alzheimer's disease (PCAD) brain. Mech Ageing Dev. 2011;132(8-9):443-8.

8. Lovell MA, Xie C, Markesbery WR. Decreased base excision repair and increased helicase activity in Alzheimer's disease brain. Brain Res. 2000;855(1):116-23.

9. Zhang J, Perry G, Smith MA, Robertson D, Olson SJ, Graham DG, et al. Parkinson's disease is associated with oxidative damage to cytoplasmic DNA and RNA in substantia nigra neurons. Am J Pathol. 1999;154(5):1423-9.

10. Alam ZI, Jenner A, Daniel SE, Lees AJ, Cairns N, Marsden CD, et al. Oxidative DNA damage in the parkinsonian brain: an apparent selective increase in 8-hydroxyguanine levels in substantia nigra. J Neurochem. 1997;69(3):1196-203.

11. Devi L, Raghavendran V, Prabhu BM, Avadhani NG, Anandatheerthavarada HK. Mitochondrial import and accumulation of alpha-synuclein impair complex I in human dopaminergic neuronal cultures and Parkinson disease brain. J Biol Chem. 2008;283(14):9089-100.

12. Sanders LH, Paul KC, Howlett EH, Lawal H, Boppana S, Bronstein JM, et al. Editor's Highlight: Base Excision Repair Variants and Pesticide Exposure Increase Parkinson's Disease Risk. Toxicol Sci. 2017;158(1):188-98.

13. Kruman, II, Kumaravel TS, Lohani A, Pedersen WA, Cutler RG, Kruman Y, et al. Folic acid deficiency and homocysteine impair DNA repair in hippocampal neurons and sensitize them to amyloid toxicity in experimental models of Alzheimer's disease. J Neurosci. 2002;22(5):1752-62.

14. Kruman, II, Schwartz E, Kruman Y, Cutler RG, Zhu X, Greig NH, et al. Suppression of uracil-DNA glycosylase induces neuronal apoptosis. J Biol Chem. 2004;279(42):43952-60.

15. Ravaglia G, Forti P, Maioli F, Martelli M, Servadei L, Brunetti N, et al. Homocysteine and folate as risk factors for dementia and Alzheimer disease. Am J Clin Nutr. 2005;82(3):636-43.

16. Chan MK, Ocampo-Hafalla MT, Vartanian V, Jaruga P, Kirkali G, Koenig KL, et al. Targeted deletion of the genes encoding NTH1 and NEIL1 DNA N-glycosylases reveals the existence of novel carcinogenic oxidative damage to DNA. DNA Repair (Amst). 2009;8(7):786-94.

17. Canugovi C, Shamanna RA, Croteau DL, Bohr VA. Base excision DNA repair levels in mitochondrial lysates of Alzheimer's disease. Neurobiol Aging. 2014;35(6):1293-300.

18. Rolseth V, Runden-Pran E, Luna L, McMurray C, Bjoras M, Ottersen OP. Widespread distribution of DNA glycosylases removing oxidative DNA lesions in human and rodent brains. DNA Repair (Amst). 2008;7(9):1578-88.

19. Canugovi C, Misiak M, Scheibye-Knudsen M, Croteau DL, Mattson MP, Bohr VA. Loss of NEIL1 causes defects in olfactory function in mice. Neurobiol Aging. 2015;36(2):1007-12.

20. Sugo N, Aratani Y, Nagashima Y, Kubota Y, Koyama H. Neonatal lethality with abnormal neurogenesis in mice deficient in DNA polymerase beta. EMBO J. 2000;19(6):1397-404.

21. Sykora P, Misiak M, Wang Y, Ghosh S, Leandro GS, Liu D, et al. DNA polymerase beta deficiency leads to neurodegeneration and exacerbates Alzheimer disease phenotypes. Nucleic Acids Res. 2015;43(2):943-59.

22. Wang H, Chen Y, Chen J, Zhang Z, Lao W, Li X, et al. Cell cycle regulation of DNA polymerase beta in rotenone-based Parkinson's disease models. PLoS One. 2014;9(10):e109697.

23. Fymat A. Dementia: A Review. J Clin Psychiatr Neurosci. 2018;Vol.1(No.3):27-34.

24. Bakanlığı AvSP. Demans Bakım Modeli Raporu. 2017.

25. Ripich DN, Horner J. The Neurodegenerative Dementias: Diagnoses and Interventions. The ASHA Leader. 2004.

26. Association As. 2020 Alzheimer’s Disease Facts and Figures. 2020.

27. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res. 1975;12(3):189-98.

28. Hughes CP, Berg L, Danziger WL, Coben LA, Martin RL. A new clinical scale for the staging of dementia. Br J Psychiatry. 1982;140:566-72.

29. Masters CL, Bateman R, Blennow K, Rowe CC, Sperling RA, Cummings JL. Alzheimer's disease. Nat Rev Dis Primers. 2015;1:15056.

30. Grober E, Dickson D, Sliwinski MJ, Buschke H, Katz M, Crystal H, et al. Memory and mental status correlates of modified Braak staging. Neurobiol Aging. 1999;20(6):573-9.

31. Hernandez-Zimbron LF, Rivas-Arancibia S. Deciphering an interplay of proteins associated with amyloid beta 1-42 peptide and molecular mechanisms of Alzheimer's disease. Rev Neurosci. 2014;25(6):773-83.

32. Golde TE, Eckman CB, Younkin SG. Biochemical detection of Abeta isoforms: implications for pathogenesis, diagnosis, and treatment of Alzheimer's disease. Biochim Biophys Acta. 2000;1502(1):172-87.

33. Selkoe DJ. Alzheimer's disease: genes, proteins, and therapy. Physiol Rev. 2001;81(2):741-66.

34. Bekris LM, Yu CE, Bird TD, Tsuang DW. Genetics of Alzheimer disease. J Geriatr Psychiatry Neurol. 2010;23(4):213-27.

35. Goldman JS, Hahn SE, Catania JW, LaRusse-Eckert S, Butson MB, Rumbaugh M, et al. ADDENDUM: Genetic counseling and testing for Alzheimer disease: joint practice guidelines of the American College of Medical Genetics and the National Society of Genetic Counselors. Genet Med. 2019;21(10):2404.

36. Lott IT, Dierssen M. Cognitive deficits and associated neurological complications in individuals with Down's syndrome. Lancet Neurol. 2010;9(6):623-33.

37. Green RC, Cupples LA, Go R, Benke KS, Edeki T, Griffith PA, et al. Risk of dementia among white and African American relatives of patients with Alzheimer disease. JAMA. 2002;287(3):329-36.

38. Liu CC, Liu CC, Kanekiyo T, Xu H, Bu G. Apolipoprotein E and Alzheimer disease: risk, mechanisms and therapy. Nat Rev Neurol. 2013;9(2):106-18.

39. Ungar L, Altmann A, Greicius MD. Apolipoprotein E, gender, and Alzheimer's disease: an overlooked, but potent and promising interaction. Brain Imaging Behav. 2014;8(2):262-73.

40. Van Cauwenberghe C, Van Broeckhoven C, Sleegers K. The genetic landscape of Alzheimer disease: clinical implications and perspectives. Genet Med. 2016;18(5):421-30.

41. Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet. 2013;45(12):1452-8.

42. Harold D, Abraham R, Hollingworth P, Sims R, Gerrish A, Hamshere ML, et al. Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. Nat Genet. 2009;41(10):1088-93.

43. Lambert JC, Heath S, Even G, Campion D, Sleegers K, Hiltunen M, et al. Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. Nat Genet. 2009;41(10):1094-9.

44. Seshadri S, Fitzpatrick AL, Ikram MA, DeStefano AL, Gudnason V, Boada M, et al. Genome-wide analysis of genetic loci associated with Alzheimer disease. JAMA. 2010;303(18):1832-40.

45. Naj AC, Jun G, Beecham GW, Wang LS, Vardarajan BN, Buros J, et al. Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. Nat Genet. 2011;43(5):436-41.

46. Hollingworth P, Harold D, Sims R, Gerrish A, Lambert JC, Carrasquillo MM, et al. Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease. Nat Genet. 2011;43(5):429-35.

47. Xu W, Tan L, Wang HF, Jiang T, Tan MS, Tan L, et al. Meta-analysis of modifiable risk factors for Alzheimer's disease. J Neurol Neurosurg Psychiatry. 2015;86(12):1299-306.

48. Iacono D, Resnick SM, O'Brien R, Zonderman AB, An Y, Pletnikova O, et al. Mild cognitive impairment and asymptomatic Alzheimer disease subjects: equivalent beta-amyloid and tau loads with divergent cognitive outcomes. J Neuropathol Exp Neurol. 2014;73(4):295-304.

49. Maarouf CL, Daugs ID, Kokjohn TA, Walker DG, Hunter JM, Kruchowsky JC, et al. Alzheimer's disease and non-demented high pathology control nonagenarians: comparing and contrasting the biochemistry of cognitively successful aging. PLoS One. 2011;6(11):e27291.

50. Soltys DT, Pereira CPM, Rowies FT, Farfel JM, Grinberg LT, Suemoto CK, et al. Lower mitochondrial DNA content but not increased mutagenesis associates with decreased base excision repair activity in brains of AD subjects. Neurobiol Aging. 2019;73:161-70.

51. Hoehn MM, Yahr MD. Parkinsonism: onset, progression and mortality. Neurology. 1967;17(5):427-42.

52. Goetz CG, Tilley BC, Shaftman SR, Stebbins GT, Fahn S, Martinez-Martin P, et al. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. Mov Disord. 2008;23(15):2129-70.

53. Schwab, R.S. and England, A.C. Projection Technique for Evaluating Surgery in Parkinson's Disease. In: Billingham, F.H. and Donaldson, M.C., Eds., Third Symposium on Parkinson's Disease, Churchill Livingstone, Edinburgh, 152-157

54. Kalia LV, Lang AE. Parkinson's disease. The Lancet 2015;386:896-912.

55. Blauwendraat C, Nalls MA, Singleton AB. The genetic architecture of Parkinson's disease. Lancet Neurol. 2020;19(2):170-8.

56. Sykora P, Wilson DM, 3rd, Bohr VA. Base excision repair in the mammalian brain: implication for age related neurodegeneration. Mech Ageing Dev. 2013;134(10):440-8.

57. Maynard S, Schurman SH, Harboe C, de Souza-Pinto NC, Bohr VA. Base excision repair of oxidative DNA damage and association with cancer and aging. Carcinogenesis. 2009;30(1):2-10.

58. Kim YJ, Wilson DM, 3rd. Overview of base excision repair biochemistry. Curr Mol Pharmacol. 2012;5(1):3-13.

59. Nickson CM, Parsons JL. Monitoring regulation of DNA repair activities of cultured cells in-gel using the comet assay. Front Genet. 2014;5:232.

60. Petermann E, Ziegler M, Oei SL. ATP-dependent selection between single nucleotide and long patch base excision repair. DNA Repair (Amst). 2003;2(10):1101-14.

61. Fortini P, Dogliotti E. Base damage and single-strand break repair: mechanisms and functional significance of short- and long-patch repair subpathways. DNA Repair (Amst). 2007;6(4):398-409.

62. Sung JS, Demple B. Roles of base excision repair subpathways in correcting oxidized abasic sites in DNA. FEBS J. 2006;273(8):1620-9.

63. Kim HS, Guo C, Thompson EL, Jiang Y, Kelley MR, Vasko MR, et al. APE1, the DNA base excision repair protein, regulates the removal of platinum adducts in sensory neuronal cultures by NER. Mutat Res. 2015;779:96-104.

64. Yang JL, Sykora P, Wilson DM, 3rd, Mattson MP, Bohr VA. The excitatory neurotransmitter glutamate stimulates DNA repair to increase neuronal resiliency. Mech Ageing Dev. 2011;132(8-9):405-11.

65. Yu H, Su Y, Shin J, Zhong C, Guo JU, Weng YL, et al. Tet3 regulates synaptic transmission and homeostatic plasticity via DNA oxidation and repair. Nat Neurosci. 2015;18(6):836-43.

66. Hou Y, Song H, Croteau DL, Akbari M, Bohr VA. Genome instability in Alzheimer disease. Mech Ageing Dev. 2017;161(Pt A):83-94.

67. Shao C, Xiong S, Li GM, Gu L, Mao G, Markesbery WR, et al. Altered 8-oxoguanine glycosylase in mild cognitive impairment and late-stage Alzheimer's disease brain. Free Radic Biol Med. 2008;45(6):813-9.

68. Canugovi C, Yoon JS, Feldman NH, Croteau DL, Mattson MP, Bohr VA. Endonuclease VIII-like 1 (NEIL1) promotes short-term spatial memory retention and protects from ischemic stroke-induced brain dysfunction and death in mice. Proc Natl Acad Sci U S A. 2012;109(37):14948-53.

69. Lillenes MS, Rabano A, Stoen M, Riaz T, Misaghian D, Mollersen L, et al. Altered DNA base excision repair profile in brain tissue and blood in Alzheimer's disease. Mol Brain. 2016;9(1):61.

70. Lillenes MS, Stoen M, Gunther CC, Selnes P, Stenset VT, Espeseth T, et al. Mitochondrial transcription factor A (TFAM) rs1937 and AP endonuclease 1 (APE1) rs1130409 alleles are associated with reduced cognitive performance. Neurosci Lett. 2017;645:46-52.

71. Copani A, Hoozemans JJ, Caraci F, Calafiore M, Van Haastert ES, Veerhuis R, et al. DNA polymerase-beta is expressed early in neurons of Alzheimer's disease brain and is loaded into DNA replication forks in neurons challenged with beta-amyloid. J Neurosci. 2006;26(43):10949-57.

72. Kwiatkowski D, Czarny P, Toma M, Jurkowska N, Sliwinska A, Drzewoski J, et al. Associations between DNA Damage, DNA Base Excision Repair Gene Variability and Alzheimer's Disease Risk. Dement Geriatr Cogn Disord. 2016;41(3-4):152-71.

73. Kwiatkowski D, Czarny P, Toma M, Korycinska A, Sowinska K, Galecki P, et al. Association between Single-Nucleotide Polymorphisms of the hOGG1,NEIL1,APEX1, FEN1,LIG1, and LIG3 Genes and Alzheimer's Disease Risk. Neuropsychobiology. 2016;73(2):98-107.

74. Parildar-Karpuzoglu H, Dogru-Abbasoglu S, Hanagasi HA, Karadag B, Gurvit H, Emre M, et al. Single nucleotide polymorphisms in base-excision repair genes hOGG1, APE1 and XRCC1 do not alter risk of Alzheimer's disease. Neurosci Lett. 2008;442(3):287-91.

75. Coppede F, Mancuso M, Lo Gerfo A, Manca ML, Petrozzi L, Migliore L, et al. A Ser326Cys polymorphism in the DNA repair gene hOGG1 is not associated with sporadic Alzheimer's disease. Neurosci Lett. 2007;414(3):282-5.

76. Dorszewska J, Kempisty B, Jaroszewska-Kolecka J, Rozycka A, Florczak J, Lianeri M, et al. Expression and polymorphisms of gene 8-oxoguanine glycosylase 1 and the level of oxidative DNA damage in peripheral blood lymphocytes of patients with Alzheimer's disease. DNA Cell Biol. 2009;28(11):579-88.

77. Mao G, Pan X, Zhu BB, Zhang Y, Yuan F, Huang J, et al. Identification and characterization of OGG1 mutations in patients with Alzheimer's disease. Nucleic Acids Res. 2007;35(8):2759-66.

78. Kwiatkowski D, Czarny P, Galecki P, Bachurska A, Talarowska M, Orzechowska A, et al. Variants of Base Excision Repair Genes MUTYH , PARP1 and XRCC1 in Alzheimer's Disease Risk. Neuropsychobiology. 2015;71(3):176-86.

79. Dogru-Abbasoglu S, Aykac-Toker G, Hanagasi HA, Gurvit H, Emre M, Uysal M. The Arg194Trp polymorphism in DNA repair gene XRCC1 and the risk for sporadic late-onset Alzheimer's disease. Neurol Sci. 2007;28(1):31-4.

80. Qian Y, Chen W, Wu J, Tao T, Bi L, Xu W, et al. Association of polymorphism of DNA repair gene XRCC1 with sporadic late-onset Alzheimer's disease and age of onset in elderly Han Chinese. J Neurol Sci. 2010;295(1-2):62-5.

81. Liu HP, Lin WY, Wu BT, Liu SH, Wang WF, Tsai CH, et al. Evaluation of the poly(ADP-ribose) polymerase-1 gene variants in Alzheimer's disease. J Clin Lab Anal. 2010;24(3):182-6.

82. Scott TL, Wicker CA, Suganya R, Dhar B, Pittman T, Horbinski C, et al. Polyubiquitination of apurinic/apyrimidinic endonuclease 1 by Parkin. Mol Carcinog. 2017;56(2):325-36.

83. Fukae J, Takanashi M, Kubo S, Nishioka K, Nakabeppu Y, Mori H, et al. Expression of 8-oxoguanine DNA glycosylase (OGG1) in Parkinson's disease and related neurodegenerative disorders. Acta Neuropathol. 2005;109(3):256-62.

84. Kam TI, Mao X, Park H, Chou SC, Karuppagounder SS, Umanah GE, et al. Poly(ADP-ribose) drives pathologic alpha-synuclein neurodegeneration in Parkinson's disease. Science. 2018;362(6414).

85. Xue B, Xiao W, Tian H. Nei-like 1 inhibition results in motor dysfunction and promotes inflammation in Parkinson's disease mice model. Biochem Biophys Res Commun. 2020;521(1):245-51.

86. Coppede F, Ceravolo R, Migheli F, Fanucchi F, Frosini D, Siciliano G, et al. The hOGG1 Ser326Cys polymorphism is not associated with sporadic Parkinson's disease. Neurosci Lett. 2010;473(3):248-51.

87. Gencer M, Dasdemir S, Cakmakoglu B, Cetinkaya Y, Varlibas F, Tireli H, et al. DNA repair genes in Parkinson's disease. Genet Test Mol Biomarkers. 2012;16(6):504-7.

88. KALEAĞASI H, EDGÜNLÜ TG, ERDAL ME, Okan, DOĞU. The ARG399GLN Polymorphism in DNA Repair Gene XRCC1 Does Not Alter Risk of Parkinson's Disease. Journal of Neurological Sciences ·. 2009:185-9.

89. Brighina L, Riva C, Bertola F, Fermi S, Saracchi E, Piolti R, et al. Association analysis of PARP1 polymorphisms with Parkinson's disease. Parkinsonism Relat Disord. 2011;17(9):701-4.

90. Luoma PT, Eerola J, Ahola S, Hakonen AH, Hellstrom O, Kivisto KT, et al. Mitochondrial DNA polymerase gamma variants in idiopathic sporadic Parkinson disease. Neurology. 2007;69(11):1152-9.

91. Beard WA, Wilson SH. Structure and mechanism of DNA polymerase Beta. Chem Rev. 2006;106(2):361-82.

92. Beard WA, Wilson SH. Structure and mechanism of DNA polymerase beta. Biochemistry. 2014;53(17):2768-80.

93. Su AI, Wiltshire T, Batalov S, Lapp H, Ching KA, Block D, et al. A gene atlas of the mouse and human protein-encoding transcriptomes. Proc Natl Acad Sci U S A. 2004;101(16):6062-7.

94. Wei W, Englander EW. DNA polymerase beta-catalyzed-PCNA independent long patch base excision repair synthesis: a mechanism for repair of oxidatively damaged DNA ends in post-mitotic brain. J Neurochem. 2008;107(3):734-44.

95. Slupphaug G, Markussen FH, Olsen LC, Aasland R, Aarsaether N, Bakke O, et al. Nuclear and mitochondrial forms of human uracil-DNA glycosylase are encoded by the same gene. Nucleic Acids Res. 1993;21(11):2579-84.

96. Haug T, Skorpen F, Kvaloy K, Eftedal I, Lund H, Krokan HE. Human uracil-DNA glycosylase gene: sequence organization, methylation pattern, and mapping to chromosome 12q23-q24.1. Genomics. 1996;36(3):408-16.

97. Nilsen H, Otterlei M, Haug T, Solum K, Nagelhus TA, Skorpen F, et al. Nuclear and mitochondrial uracil-DNA glycosylases are generated by alternative splicing and transcription from different positions in the UNG gene. Nucleic Acids Res. 1997;25(4):750-5.

98. Forsbring M, Vik ES, Dalhus B, Karlsen TH, Bergquist A, Schrumpf E, et al. Catalytically impaired hMYH and NEIL1 mutant proteins identified in patients with primary sclerosing cholangitis and cholangiocarcinoma. Carcinogenesis. 2009;30(7):1147-54.

99. Wu L, Zhao L. ApoE2 and Alzheimer's disease: time to take a closer look. Neural Regen Res. 2016;11(3):412-3.

100. 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. 1993;261(5123):921-3.

101. Harhangi BS, de Rijk MC, van Duijn CM, Van Broeckhoven C, Hofman A, Breteler MM. APOE and the risk of PD with or without dementia in a population-based study. Neurology. 2000;54(6):1272-6.

102. Li YJ, Hauser MA, Scott WK, Martin ER, Booze MW, Qin XJ, et al. Apolipoprotein E controls the risk and age at onset of Parkinson disease. Neurology. 2004;62(11):2005-9.

103. Barzon L, Lavezzo E, Costanzi G, Franchin E, Toppo S, Palu G. Next-generation sequencing technologies in diagnostic virology. J Clin Virol. 2013;58(2):346-50.

104. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative Genomics Viewer. Nature Biotechnology. 2011;29:24-6.

105. Clarke GM, Anderson CA, Pettersson FH, Cardon LR, Morris AP, Zondervan KT. Basic statistical analysis in genetic case-control studies. Nat Protoc. 2011;6(2):121-33.

106. Kvaloy K, Nilsen H, Steinsbekk KS, Nedal A, Monterotti B, Akbari M, et al. Sequence variation in the human uracil-DNA glycosylase (UNG) gene. Mutat Res. 2001;461(4):325-38.

107. Lo SF, Wan L, Huang CM, Lin HC, Chen SY, Liu SC, et al. Genetic polymorphisms of the DNA repair gene UNG are associated with the susceptibility of rheumatoid arthritis. Rheumatol Int. 2012;32(12):3723-7.

108. Doherty JA, Sakoda LC, Loomis MM, Barnett MJ, Julianto L, Thornquist MD, et al. DNA repair genotype and lung cancer risk in the beta-carotene and retinol efficacy trial. Int J Mol Epidemiol Genet. 2013;4(1):11-34.

109. Yin J, Sang Y, Zheng L, Wang L, Yuan L, Liu C, et al. Uracil-DNA glycosylase (UNG) rs246079 G/A polymorphism is associated with decreased risk of esophageal cancer in a Chinese population. Med Oncol. 2014;31(11):272.

110. Isbir T, Agachan B, Yilmaz H, Aydin M, Kara I, Eker E, et al. Apolipoprotein-E gene polymorphism and lipid profiles in Alzheimer's disease. Am J Alzheimers Dis Other Demen. 2001;16(2):77-81.

111. Alaylioglu M, Gezen-Ak D, Dursun E, Bilgic B, Hanagasi H, Ertan T, et al. The Association Between Clusterin and APOE Polymorphisms and Late-Onset Alzheimer Disease in a Turkish Cohort. J Geriatr Psychiatry Neurol. 2016;29(4):221-6.

112. Reiman EM, Webster JA, Myers AJ, Hardy J, Dunckley T, Zismann VL, et al. GAB2 alleles modify Alzheimer's risk in APOE epsilon4 carriers. Neuron. 2007;54(5):713-20.

113. Benitez BA, Karch CM, Cai Y, Jin SC, Cooper B, Carrell D, et al. The PSEN1, p.E318G variant increases the risk of Alzheimer's disease in APOE-epsilon4 carriers. PLoS Genet. 2013;9(8):e1003685.

114. Limon-Sztencel A, Lipska-Zietkiewicz BS, Chmara M, Wasag B, Bidzan L, Godlewska BR, et al. The algorithm for Alzheimer risk assessment based on APOE promoter polymorphisms. Alzheimers Res Ther. 2016;8(1):19.

115. Ward A, Crean S, Mercaldi CJ, Collins JM, Boyd D, Cook MN, et al. Prevalence of apolipoprotein E4 genotype and homozygotes (APOE e4/4) among patients diagnosed with Alzheimer's disease: a systematic review and meta-analysis. Neuroepidemiology. 2012;38(1):1-17.

116. Bahia VS, Kok F, Marie SN, Shinjo SO, Caramelli P, Nitrini R. Polymorphisms of APOE and LRP genes in Brazilian individuals with Alzheimer disease. Alzheimer Dis Assoc Disord. 2008;22(1):61-5.

117. de-Andrade FM, Larrandaburu M, Callegari-Jacques SM, Gastaldo G, Hutz MH. Association of apolipoprotein E polymorphism with plasma lipids and Alzheimer's disease in a Southern Brazilian population. Braz J Med Biol Res. 2000;33(5):529-37.

118. Furuya TK, Silva PN, Payao SL, Bertolucci PH, Rasmussen LT, De Labio RW, et al. Analysis of SNAP25 mRNA expression and promoter DNA methylation in brain areas of Alzheimer's Disease patients. Neuroscience. 2012;220:41-6.

119. Tamaoka A, Miyatake F, Matsuno S, Ishii K, Nagase S, Sahara N, et al. Apolipoprotein E allele-dependent antioxidant activity in brains with Alzheimer's disease. Neurology. 2000;54(12):2319-21.

120. Rebeck GW, Reiter JS, Strickland DK, Hyman BT. Apolipoprotein E in sporadic Alzheimer's disease: allelic variation and receptor interactions. Neuron. 1993;11(4):575-80.

121. Leduc V, Domenger D, De Beaumont L, Lalonde D, Belanger-Jasmin S, Poirier J. Function and comorbidities of apolipoprotein e in Alzheimer's disease. Int J Alzheimers Dis. 2011;2011:974361.

122. Laws SM, Hone E, Taddei K, Harper C, Dean B, McClean C, et al. Variation at the APOE -491 promoter locus is associated with altered brain levels of apolipoprotein E. Mol Psychiatry. 2002;7(8):886-90.

123. Nagy Z, Esiri MM, Jobst KA, Johnston C, Litchfield S, Sim E, et al. Influence of the apolipoprotein E genotype on amyloid deposition and neurofibrillary tangle formation in Alzheimer's disease. Neuroscience. 1995;69(3):757-61.

124. Yashin AI, Fang F, Kovtun M, Wu D, Duan M, Arbeev K, et al. Hidden heterogeneity in Alzheimer's disease: Insights from genetic association studies and other analyses. Exp Gerontol. 2018;107:148-60.

125. Babenko VN, Afonnikov DA, Ignatieva EV, Klimov AV, Gusev FE, Rogaev EI. Haplotype analysis of APOE intragenic SNPs. BMC Neurosci. 2018;19(Suppl 1):16.

126. Bratosiewicz-Wasik J, Liberski PP, Peplonska B, Styczynska M, Smolen-Dzirba J, Cycon M, et al. Regulatory region single nucleotide polymorphisms of the apolipoprotein E gene as risk factors for Alzheimer's disease. Neurosci Lett. 2018;684:86-90.

127. Lescai F, Chiamenti AM, Codemo A, Pirazzini C, D'Agostino G, Ruaro C, et al. An APOE haplotype associated with decreased epsilon4 expression increases the risk of late onset Alzheimer's disease. J Alzheimers Dis. 2011;24(2):235-45.

128. Bizzarro A, Seripa D, Acciarri A, Matera MG, Pilotto A, Tiziano FD, et al. The complex interaction between APOE promoter and AD: an Italian case-control study. Eur J Hum Genet. 2009;17(7):938-45.

129. Kavvoura FK, McQueen MB, Khoury MJ, Tanzi RE, Bertram L, Ioannidis JP. Evaluation of the potential excess of statistically significant findings in published genetic association studies: application to Alzheimer's disease. Am J Epidemiol. 2008;168(8):855-65.

130. Xiao H, Gao Y, Liu L, Li Y. Association between polymorphisms in the promoter region of the apolipoprotein E (APOE) gene and Alzheimer's disease: A meta-analysis. EXCLI J. 2017;16:921-38.

7. APPENDICIES

Appendix-1 Ethics Comitee Approval

ACIBADEM
ÜNİVERSİTESİ

Sayı: B.30.2.ACÜ.0.00.00.050-06/45

Konu: ATADEK 2014-700

05 Eylül 2014

Sayın Doç. Dr. Meltem Müftüoğlu;

ATADEK-2014-700 kodlu ve **"Sporadik Alzheimer hastalığı için yeni bir genetik risk faktörü olarak baz ekzisyon tamir genlerinin genetik varyantlarının taranması ve değelendirilmesi"** başlıklı araştırma projeniz Acıbadem Üniversitesi Tıbbi Araştırmalar Değerlendirme Kurulu'nun 2 Eylül 2014 tarihli toplantısında incelenmiş ve etik açıdan uygun bulunmuştu.

ATADEK 2014-700 sayılı karar ektedir.

Prof. Dr. İsmail Hakkı Ulus

Tıbbi Araştırmalar Değerlendirme Kurulu Başkanı

Kerem Aydınlar Kampüsü Kayışdağı Caddesi 32 İçerenköy 34752 Ataşehir İstanbul T 0216 500 44 44 F 0216 576 50 76
www.acibadem.edu.tr

Appendix -2 The primer sequences from Ion AmpliSeq designer software

***POLB*:chr8:42,195,472:42,229,331**

Amplicon ID	Start	End	Primer Pool	Forward Primer	Reverse Primer
AMPL7160276555	42195157	42195182	Pool 1	CCAGATTTCTCTACAGCAACTACGT	GTGAACCATGCCGAGAACAGGAAT
AMPL7160090200	42195467	42195489	Pool 2	GGTTGGCAGGTGTACTAACACG	GAAGACCTGGAGTTTTGACCTAAGATATTA
AMPL7158705911	42195748	42195774	Pool 1	GGGAAAGGATTCCAGATAAACACTGA	AGTGTTCAGAACCAGGGACTAGA
AMPL7155757227	42196060	42196083	Pool 2	CAAGTCCTGGTACCTCCTTCAAG	CTACCCTACCCACTGGACTGT
AMPL7157138981	42196230	42196255	Pool 1	CTTTCTTCTTTCCTTCCAGCCTCTT	GCTCGGGTGAACAAGAACCA
AMPL7157138982	42196459	42196482	Pool 2	CCCTTCCAGAAAACAGTTCTCGT	CAGCCTCGATTCTTGCTTTTTCC
AMPL7157138983	42196664	42196686	Pool 1	GCCCAGTGGATATTTGGTCCAT	CCTGGTGAGATGAACAAGGACA
AMPL7157138984	42196968	42196992	Pool 2	CTGATCCTGATCCTTGTTCCTTGT	TGTGGACCATTTGAGTTTAGGAGTTG
AMPL7157138985	42197112	42197134	Pool 1	GGTCTCGCTGTGGTATGATCAT	TTTGCTTTTAAGAGCTGAGGGACTT
AMPL7157138986	42197378	42197400	Pool 2	GGTGTCAGCTCTCCTTAGCCTA	GGACTGGGACTTCTTGAGTAAATGAT
AMPL7157138987	42197508	42197533	Pool 1	AAATCTTCATTCCTTGCCTGGACTT	GCATGAAACAAAGTTTTGACCATTTTGAC
AMPL7157138988	42197811	42197838	Pool 2	GGAAAATTCCACATCCGATCTCATGTA	AGTCAAGAATGGCAAGATTGCCA
AMPL7157138989	42198108	42198133	Pool 1	GCCCTTAGCCCTCTTTTCTTACTTC	CAATCAGCTCTGAACTGTGTGGA
AMPL7157138990	42198247	42198269	Pool 2	TATCCCAAGTCAGCTCCCTTCT	GTAGGGAATCACCTGCTTTTCCA
AMPL7157138991	42198563	42198588	Pool 1	TTGCTTAACGGGTATGGTATGAGTG	CACATGACTGATGATGAATCTGAAAAGAC
AMPL7157138992	42198872	42198894	Pool 2	ACTGTGTTTGCCTGCTCCTTTA	TTCCCACTGCAGTTCAGCATAA
AMPL7157138993	42199162	42199184	Pool 1	TGTGGAGGTGGGCAAGAATTAC	TGGAAGATAAAGATGCTTACAAGAAACAGT
AMPL7157138994	42199458	42199488	Pool 2	CCTTCATCATGTCTGAAGTTATGCTATTCT	CGATAATACAATGTAGAAAGCATTGATGGT
AMPL7157138995	42199658	42199680	Pool 1	CCTCCATCCAGGTCACCAAAAA	ACACACCCAAATAGAGGATGATAGTTCATA
AMPL7157138997	42199989	42200021	Pool 2	TATATGAACTATCATCCTCTATTTGGGTGTGT	GGACTTTGTTAAAGTTGTTAAGAATAAGGCCTA
AMPL7157139302	42200075	42200107	Pool 1	TACCATCTTGTCATCTTTAAGTTTCATTTGGT	CACACGCAGCCCTTGGATA
AMPL7157139303	42201039	42201070	Pool 2	AAAAAGTAACATAGGTATCCATAAGCAGATC	CTGGCCTACGTCTAACTTTTAAAGTATCT
AMPL7157139304	42201702	42201726	Pool 1	CTGTAATCCTAGCTGGTTCACACC	CTGGCCTAATTTTCTACCTCCATATGG
AMPL7157139003	42202018	42202039	Pool 2	GGTGGTCTGTCAGCTCATCTG	CAAATATCTATGCTTATGCTTTCACCAAGA
AMPL7157135855	42202169	42202199	Pool 1	GGTGAAACATCATCTCCACTAAAATACAAG	CATCAAGGCCTTAATTTTTCTTAAACTGGA
AMPL7157139005	42202500	42202526	Pool 2	CACACAAAATAAAGAGTGGAGCTGAA	GTGCCACTTGGTCACTCTTCTTAA
AMPL7157139006	42202795	42202823	Pool 1	ACATGTACTCACGTTGAATTTCAACTTG	CAAAATAATCTGTTCCCAAGCATTTCAGAT

***POLB*:chr8:42,195,472:42,229,331**

AMPL7157139007	42203048	42203076	Pool 2	GGGTCACATGCCTCGTTAATTTGTAATA	TGGTCCTTAAATTTAAGGAACTTTGAATCA

AMPL7157139008	42203321	42203354	Pool 1	CTATCTATTGACAATCAGATTGTTGTCTCATCT	GGCTCAAGGCAGTGGATCAC
AMPL7157139009	42203411	42203442	Pool 2	TTTGTGAAAGATGTCAAATTATCATGGCTTT	AGGAGGTGATAACTATCAGGACCTG
AMPL7157139010	42203735	42203756	Pool 1	CAGCCTGAGATGCCTTTTGAG	CTGCAGCAGTAAGCTGTGATTATG
AMPL7157139011	42203855	42203879	Pool 2	TAACCTTCTCAGGAGATGAGCCAT	GCAGTTATGGTCCAGTCATGGT
AMPL7157135699	42204046	42204073	Pool 1	GCATGCTCCATCACACTCAGATATTTT	GGTAATGTTATGACTCTGAAGGGCAA
AMPL7157139013	42204451	42204480	Pool 2	ACTTTAATTCTGAGTTTCTGTCTTGGTGA	AATTACTGAGGAGTAATTCCAGAAATGG
AMPL7157139014	42204716	42204748	Pool 1	TCCTTTTCACAATTATTGAAGAATCACAGAGT	CTGGAAATTTCAGATTTCTAGAATTACTGAGAA
AMPL7157139015	42204987	42205020	Pool 2	TCCTTCTTAGTAACGGAAGTAATAGATTCTTCT	AGATAGAAATGAACAAAAGATTTAGGAAGCAGA
AMPL7157139016	42205228	42205250	Pool 1	ATGAGCCTGAGCAGTTTCCTTT	TCTTTCTCCACACTGCATGTATTTGAT
AMPL7157139017	42205504	42205532	Pool 2	ACCAAAGTTGCAAATTAAACAAGACACT	AGAGCAGAGATCCTGTCTTTTCATTTT
AMPL7157139018	42205743	42205768	Pool 1	TTTTTGATCACCTGTATGCCAGTCA	CCCTAGAGTTTCTCATTCACCATGT
AMPL7157139019	42206030	42206056	Pool 2	CCCTGAGAACTTAGGCATATTCTTGG	CCAGTGAAAATAAATTCTCCCGAAGACA
AMPL7157139020	42206326	42206353	Pool 1	TTCTCAGAGTTTCCCTTTTGTCATTCA	AGTTGCTAAAAACTCATCAATCTTTTCAGC
AMPL7157139021	42206514	42206544	Pool 2	ATTTCTAATTTTCCATGTAGCCTGGAGTAG	AACGATCTTTCAGTCAAAGTTTATAATCTCT
AMPL7157135716	42207201	42207230	Pool 1	CGGCCATCTTTTAGAATTTCTATGTTTAC	CCTGGGTGACCCACACAAAATT
AMPL7157139022	42207351	42207381	Pool 2	ACTCTGTTATTCACCTATTACTCCTTAGGT	AACTAACTGACGATTGTGTCTGCTTAA
AMPL7157139023	42207565	42207587	Pool 1	CTCGAGTTAGTGGCATTGGGTA	ATGCTCAGAGAAACAAAAGAGGCCAA
AMPL7157139024	42207855	42207880	Pool 2	CCATTTCCTCCGTGTGCATCTATTT	CCATAAAGATGAGAGCACTAAAGCCTT
AMPL7157139025	42208156	42208186	Pool 1	TTTTTGACACTGGCTGAAGTAAAGATTAAC	GTACTAAGGCAGGAGAGTCACTTTAC
AMPL7157139026	42208413	42208442	Pool 2	GTGTCTTATTCTGCTATCTAGGCTAGAGT	TCAACTTTCAGTTCTCTTGCCTAGAAAA
AMPL7157134131	42208677	42208702	Pool 1	ACCACAAGGTGTTTTAGAGAGCATC	CAGCATAGAAGAAACAGAACTGAACTC
AMPL7157139027	42208763	42208787	Pool 2	GCAGAGCCAGATTTGAACCCATAT	GAAGGACACTGAAACCTAGAAAGATGAA
AMPL7157139028	42209064	42209086	Pool 1	GCACCAGCTAGATTTGTGACCT	TTCAAGAGAAGAGAAAAACTAGGATGCAT
AMPL7157139029	42209377	42209399	Pool 2	CATAGCACCTCATACCGTCACT	GGCCATTATGTGTCAGATACTTTCCTG
AMPL7157139030	42209679	42209706	Pool 1	CCCACTTACATTCTTATTCCTGCTACA	TCAACATTGGGTGAAACTTGGGA
AMPL7157139031	42209953	42209981	Pool 2	GCAGCTCTTCTTACTATTAAGACATGGT	GCCCTGGTTCTTGGTCTTCTTA
AMPL7157139032	42210255	42210288	Pool 1	CCATCATAAACATTTTTAGGTTTGTAACTAACT	AGGTTGGTTCTTGGTCTTCTTCTTTTT
AMPL7157139305	42211021	42211054	Pool 2	GGTTATTAAAGGAATTTTTGCTTTCAAAGAAAA	TAAGGTCATCCCACTCTAATAAAGTAGTGG
AMPL7157139038	42211244	42211266	Pool 1	TTTTGGAGACTGACCTGGCTTT	AATACCTCCAAGAAACCCACAGATTC
AMPL7157139039	42211547	42211576	Pool 2	GACAGCTTATATGAGAGAATTGATTGGCT	GCATGCTTCAAAAGCAATTCCTACAAG
AMPL7157139040	42211846	42211873	Pool 1	CTCCAGAGGGAATGTGTATATTTGTGG	TTCTGGGCAGACAGAGCAAGAC
AMPL7157139041	42212037	42212070	Pool 2	CAGGTACAGTAATAATTTTCCTTTCAATGGAAT	CATTAGCACTGAGTAGTAGATTTGTGCA
				***POLB*:chr8:42,195,472:42,229,331**	
AMPL7157135868	42212354	42212374	Pool 1	CCGGCTCAACAGGATTTTGA	AGAATCACACCAAATTTTACTGACTATAGTGG

AMPL7157139043	42212506	42212539	Pool 2	AAATTAAACCACTATAGTCAGTAAAATTTGGTG	GAGCTAGATAAAAGTGACTCTTCAGGG
AMPL7157135869	42212850	42212871	Pool 1	CTGGCCCTGAAGAGTCACTTT	CAAGTGTCAAAAGAAAATCTGCCATCTTA
AMPL7157139046	42213047	42213076	Pool 2	TGAAGATAAATTGAACCATCATCAGCGAA	CCCTGCTAAAGTCAAAAGAATAAGGGA
AMPL7157139047	42213257	42213287	Pool 1	TGCTTTCACTTTTGCTTTCTAGTTTACTTG	GTTCCCAGACTATATCTGGCACA
AMPL7157139048	42213552	42213582	Pool 2	CTACTAGGCATGTTTCTATATCAGCATCAA	CACCTCCATTTACTACAGAAATGCCA
AMPL7157139049	42213782	42213804	Pool 1	TGAGAGACCTGGATGGTATGGT	GGGTTCTTTCTGAGATAACAAAAATGTCCT
AMPL7157139050	42214025	42214048	Pool 2	GCACAGTGTAGGAACCATCACTA	AGTTTACCGCTCTTATATCCTGCAAAATTA
AMPL7157139051	42214260	42214288	Pool 1	GTTGTCTGAGTTTATGTTGCAGGATAGT	CAGATCATTCTCACAAATCAAGGCTAG
AMPL7157135758	42214528	42214552	Pool 2	CCCAATTTTGCTGTTGTCATCTCA	ACCAAGTGACACTCTCAATTCTAATCAAAG
AMPL7157139052	42214724	42214753	Pool 1	GAGATGTTACAAATGCAAGTAAGATGTGT	GCAAAAATAACCCAAGATTAGGAAGTATGT
AMPL7157139053	42214881	42214903	Pool 2	TTGCTACAGTCTGTGGCAGTTT	GGAAAAAGAAAGGTAAATTAACAAAACTACCCA
AMPL7157139054	42215187	42215211	Pool 1	CCTTTTGGCTGGATGATAGTGAAG	TCTTATAGAAAACTTGGGCAGTTGGG
AMPL7157135871	42215535	42215557	Pool 2	GCCCAACTGCCCAAGTTTTCTA	GCGCCTGGCCAAAAGTTAATATTTT
AMPL7157136570	42216086	42216119	Pool 1	TGTTAGAGTAGATAGAGTAGATACCTGAACTGA	CTCCAACTCTAGATTCTTACAGCAGC
AMPL7157139307	42216151	42216174	Pool 2	AGTAGGTGATGGCCTATCAGGAA	GCTGTGCTTCCAAATCTACATTTTGAAAA
AMPL7157139309	42216430	42216458	Pool 1	CAAATAAGAATCTCCTTAGGTGGGAACT	ATGTAACTTTATGAAAGAAAGGGAGAGAAG
AMPL7157139310	42216736	42216769	Pool 2	AGTATACCATCTTTATTTCTTTCTTCTCTCCCT	CCCAGGCAACATAGTAAAACCATG
AMPL7157139311	42217058	42217086	Pool 1	TTTTTGGTAGGGACATGGTTTTACTATG	GCGGACTGGATAAACAAATGAAGTATG
AMPL7157139313	42217191	42217218	Pool 2	TTTTTGTAGCTGTGTAACATTCAGTTG	AAAAAGAAGACTAGACTGAATGTGTACTG
AMPL7157139064	42217427	42217451	Pool 1	GTTTCCATGGATGTCCAGTACACA	CCAAATATTTAAATATAAAGAAGGCCAGGCC
AMPL7157139314	42217770	42217792	Pool 2	CGGCCTGGCCTTCTTTATATTT	CATATGTAAAAGCATCCACTATAGTGCCT
AMPL7157139067	42217928	42217957	Pool 1	GAAAGGGAGGAGATAGATTGGATTCTTTT	GAGAGATCCAGGCACTCTCTTTTT
AMPL7157139068	42218229	42218259	Pool 2	TTGAGTTCAACCATAAGAATTGAAATGAGT	GGCATTAGATATACATTGCATACAACGAGT
AMPL7157139070	42218566	42218595	Pool 1	TCGTTGTATGCAATGTATATCTAATGCCA	GACAGAAGACACATATGAGATTGTCCAT
AMPL7157139071	42218745	42218767	Pool 2	GCTTGGTTCCATTTGCCCAATT	CAGCAACTCATGGAAGAATAATAGGTATCC
AMPL7157139072	42218997	42219026	Pool 1	TCTGAAGAGCTTTGTACTGATTGAATTCT	TGTTTCACCTTCCAAAGGCATATATTCTAA
AMPL7157139073	42219219	42219248	Pool 2	CTTCTATTAGAAAACCTGTGAAGGCAAAA	CGGCCATCTCTATGTTTTCTAATGT
AMPL7157139315	42219840	42219871	Pool 1	GAAAACTCATGTTAAAAATGTTTTCTTCCCG	TTAGAGACACAGTCTTGCTATATTGCAC
AMPL7157139316	42219991	42220017	Pool 2	CCATGATTGTACCTGTGAATAGCACA	CCCTGCTATTATTTTTCTTTAAAGTCTCATG
AMPL7157139317	42220305	42220336	Pool 1	GAATTACAGTCACCAAATAGAGTATCCATGA	GGCATCCTGCTGATTATTTTTCTTTCTTTT
AMPL7157136859	42220648	42220678	Pool 2	AAAAAGAAAGAAAAATAATCAGCAGGATGC	CCATGCCCAGGCCTAATTAACT
AMPL7157139077	42220729	42220757	Pool 1	TTCTCCACTTGTAAATAACACGTGTCAT	AGCAGATTATGCCTAATTAACCCATTCTTT
				***POLB*:chr8:42,195,472:42,229,331**	
AMPL7157139322	42220818	42220839	Pool 2	CAGCGGATTACAAGGTCAGGT	AAAAACTTACTGTGTGTCAAAGTCATTACTG

AMPL7157139326	42221391	42221419	Pool 1	TGATTCTGTATAAACACCAACAGGAAGG	ATCTTAATTCTAGGCTGGGCACG
AMPL7157139327	42221397	42221421	Pool 2	TGTATAAACACCAACAGGAAGGCA	AAAAAGGAACTAAGAATCTTAATTCTAGGCT
AMPL7157139084	42222043	42222066	Pool 1	GGGCCTCTTCTTTCTTAAGGTGA	ACTAATTTTAGCACATCCACTCTCCAAA
AMPL7157139085	42222128	42222152	Pool 2	CCGGGTGATATGACTGAAAGGAAA	CCACTGTGCCTACTATTTTAACCATTTTT
AMPL7157139086	42222314	42222341	Pool 1	GGGTAATAGGGAGTTCCTGTTTAAATG	TTGTATGTGTGTGGGTTAAATATACACA
AMPL7157139087	42222617	42222648	Pool 2	AAAAATAGATGGCCAATAGATGGTAAATTGT	AAAGTTCGTTTGGAAGCATTGTTAAAGT
AMPL7157139088	42222712	42222739	Pool 1	TTAGTGGGCTGTGAAAGTGTATTTTCT	CCAGACGTAGCAAGATACTTTTCTACTG
AMPL7157139089	42223018	42223048	Pool 2	TCCTAAAAACAGTATCTGCTAGACAAGTTG	GCCTGGCCCTTTTACCTTGTATT
AMPL7157139090	42223286	42223313	Pool 1	AAACCAACAGAAGCTATAGGAATCCAG	CCAGCCCACTAGATTATACTTTTTCCTTTT
AMPL7157139092	42223636	42223665	Pool 2	GAAAAAGTATAATCTAGTGGGCTGGATGT	CTGAGATGGATTCCTATGTCGCAA
AMPL7157139093	42223939	42223969	Pool 1	AGATTCAGGTGACAATAAGAACAAGAAACT	ACTCAGGCCTTTTAGAAGTAAATCCAC
AMPL7157139094	42224232	42224259	Pool 2	GCTCCATTACTAATACTCTGTATGGCA	TGCTCACTCTCCTCATTCTGTTCT
AMPL7157139095	42224544	42224570	Pool 1	GAAGGGATAGAAAGAGAGAGGATGGT	CCCTAAGTCTCATGTTCTCAATGAGG
AMPL7157139096	42224846	42224872	Pool 2	TGCAGGAGGCAGTATTAAATAGTGTG	ATGGAATGAGAATGGTTTATTTCACTACCT
AMPL7157139097	42225029	42225057	Pool 1	GCTTCTAAAAGGAAATTGGCTGCAAAAT	CAAAATACTGAAATCTGAATGGCACTCA
AMPL7157139098	42225327	42225353	Pool 2	AGTTTCCTCTGCATTCAACTATTGGT	TCATGGCTTTGTTCCTTTCCTCAA
AMPL7157139099	42225600	42225624	Pool 1	GCACTTCAGAAAAAGTGCAGGAAA	CCTCAAGAGTGGTCAAGTTAACACA
AMPL7157139100	42225904	42225934	Pool 2	TTTCTATTTTGCTGAGTGAAAAAGTATGGC	CTAGAAGGGAGAAAGACACGGAAAATAAGA
AMPL7157139101	42226171	42226199	Pool 1	AATTACCTGTGAAGGAAGAAAGTGATCC	ACAGACATCAAAATAAACAGAAGAGAGCAA
AMPL7157139102	42226431	42226460	Pool 2	TTTGTTTTCTTCTTCCTGCATAATCATCG	AGCTCTAGAATTCTGAGCACTTTTAACTT
AMPL7157139103	42226621	42226650	Pool 1	ATAGAAAGGTAGGGATAGTGTATTGCTCA	AAGAACATATGGCTCTTGGGAGTAAAAA
AMPL7157139104	42226916	42226946	Pool 2	CCTGTTTGTATCTTGGAGTTCACATTCATA	CAAAATACCAAAGATAAGCCTAAGTGGTCT
AMPL7157139105	42227217	42227243	Pool 1	GCAGTCTACCTCATGATAGTCTTCTC	GAAGGTTTGCTAAACCAAAGTATCTTTGAA
AMPL7157139106	42227519	42227549	Pool 2	GTGTGTATTAGAGATCATCTCTCATCTGGA	GTGGCACTGAAAACAATGATACTTGAG
AMPL7157139107	42227799	42227828	Pool 1	TTTTGCTGTTGTTTAATGTACCTCTAGGA	GCAAGCCAAAGGAATTAGAATCTGAATTTA
AMPL7157139108	42228103	42228125	Pool 2	TCAGTGTGCATACCAGCAACAT	CAACTCGGTTCTTGGACTAGCA
AMPL7157135895	42228414	42228441	Pool 1	CCATGGTTGGATAAGCTTGTTCTAGAT	AAAAAGACACTGACTACCTCTATCCT
AMPL7157139333	42228619	42228640	Pool 2	CTCCTGGGAGAGGATAGAGGT	AAAAAGAAAATGGTAACCAAGTCACTAAAAA
AMPL7157139335	42228950	42228981	Pool 1	TTTTTAGTGACTTGGTTACCATTTTCTTTTT	GGATACAGGCCTCATTCGCTC
AMPL7157139113	42229133	42229155	Pool 2	CATCCAGTGGAAATACCGGGAA	GGTAATGTGGTTTATAGCATAGAGCAGAAT

UNG:chr12:109,534,879:109,548,798

Amplicon ID	Start	End	Primer Pool	Forward Primer	Reverse Primer
AMPL7160090178	109534875	109535223	Pool 1	GTTTGTTTTTGAGATGCCTCGGATT	CTGCTCTTTGGACAGGCTCTTA
AMPL7155382774	109535168	109535520	Pool 2	GCCTGACCAGTCTTCTCTTCTTG	GGAGAAAAAGGAGTAGAGCGTCTT
AMPL7157137438	109535469	109535751	Pool 1	CCTCCTCAGCTCCAGGATGA	ACACGTGGAGGCTATTTGGAAA
AMPL7157137439	109535699	109536073	Pool 2	GCGCCTCTGACTCGGTAAAC	GGAATTGGGAATTAGGTTCTAAACTGGAG
AMPL7157119888	109535786	109536146	Pool 1	ATGCTAAAGGGCCAGCCAAT	GCAGAAGACGCCCATTTGTG
AMPL7155381294	109536094	109536444	Pool 2	GCTCTTACTGTCCGCTTTTGCT	CCTTGATAAAATACGGTTTCCCGAACT
AMPL7157137440	109536382	109536744	Pool 1	GCTTTGGAGAGAGCTGGAAGAAG	CTCAAGCCAGGTTTCATTCAGGA
AMPL7157137441	109536689	109537044	Pool 2	CAGGAAAGGCTGGCGTAGTAAA	GGGTGGATAAACAGTGTAATGCTTTC
AMPL7157135503	109536981	109537218	Pool 1	ATCTTTAAATCAGCTAATGGGATTTGTTGC	CAAACCTCGGGATTTCAGTGTTTTATTT
AMPL7157137442	109537158	109537414	Pool 2	CTGGAGTTAAGCCACAGTCCAT	AAAAGATTTAACTGCTTTGTGTGTCAGTT
AMPL7157137443	109537353	109537698	Pool 1	CGGTGTCAGATGTTATGGTGGT	ACCAGGTTTCATTGGTCTTGAACTC
AMPL7157137444	109537566	109537940	Pool 2	GGCTCTCATCCATCTTGTGTTCA	CGTGGCCCTATTTAAAGAGGGTT
AMPL7157137445	109537883	109538162	Pool 1	CGTGGCTTCTCACTGTTGAACTAT	GGGTAAGAAACCTAAAAGTCCAGTGT
AMPL7157137446	109538104	109538471	Pool 2	TCTGAGCTGCTACGATGCATTT	GTTTTTGAGCATTCCCTCAAAATATTCCTT
AMPL7157137491	109538396	109538655	Pool 1	TGCCCTATTCTGAATACCAGTAATCAGTA	TGACATACTCTATCTTTCAGAATCCAGCT
AMPL7157137492	109538596	109538916	Pool 2	CCCAAAAGCCTCCTGGTCAT	AAAAAGCAATTCAGATTTCCCACCA
AMPL7157137493	109538867	109539228	Pool 1	GGGAGCAGACACCCTCATA	CCCAGTTACTTTTCGATGAATTAAAGC
AMPL7157137451	109539184	109539554	Pool 2	CATGCCAGGCCCTAAATTGC	ACTTGTAAAGATATGTACCTAGCTGCTTTTT
AMPL7157137452	109539328	109539649	Pool 1	TAAGGGAGGCCGTGTGTCTAC	CTGCCTCAAACAAACATAAACAAAACAAAT
AMPL7157137453	109539581	109539786	Pool 2	CAACTCTGCTATTAGCAATCTATGCCAT	GGCCTTTGAACACTAAAGCAGAG
AMPL7157137454	109539726	109540084	Pool 1	GACAGGATCCATATCATGGACCTAATC	GGCAATGTCTGCACATTTTTCATTATCA
AMPL7157137455	109540023	109540357	Pool 2	CTACCACTGGAAACCAGGAGCAA	GGACAGTTTTCTGTCCTTAGCATATTTATG
AMPL7157137494	109540295	109540460	Pool 1	CCACTGCTTTTGGAGGATTTGG	GCCAAAGACCTTTACTTTAGTCTGTACT
AMPL7156710119	109540394	109540757	Pool 2	GTTGATTTACTTAGGTTTTCCAGAGTGC	AAAAATCTAGCAGTCGCTGGC
AMPL7157137495	109540719	109541090	Pool 1	GTGGGCCAAGCAAGGTAAGCC	AGCAGCCACTACCATCAACAAT
AMPL7157137459	109541060	109541250	Pool 2	CCCGACTCCATTGTTGATGG	AGAAGGAGAACACCTATACACATGAGAG
AMPL7157137460	109541190	109541562	Pool 1	GCTGATTTGCCTGAGCCTACAT	CCCTTATAACAGGCCTAAGGGAAAATAAAA
AMPL7157137461	109541498	109541807	Pool 2	TGTTTGGTCCTGGAAAATCCATGT	TTTGTGATTGAGTGGAATTCTTTGTGTTTT
AMPL7157137462	109541742	109542116	Pool 1	ATGTTTCTGGCTTCCAAGAAATTTG	CTTTGGAGGTTGAGGAGGACAG
AMPL7157137463	109541833	109542205	Pool 2	CAGTGATCATTATTGTTTTGTTTCTGGGT	GGTCGTGTTGGGTATCTGCAAT
AMPL7157137464	109542148	109542469	Pool 1	CGGCCTATCGGTGATCATTATTAACC	AGCCTCCATCTCTAGTCCATCAC
AMPL7157137465	109542412	109542728	Pool 2	GGCATGAGTCACCATCATCTTTCT	AGTGCTCAACCTAGTACCCACA

UNG:chr12:109,534,879:109,548,798					
AMPL7157137466	109542667	109543041	Pool 1	GGGTATCAGAAATGTGCACGTATATTTAAC	TCTGCATGCAAAGTACTCAACCT
AMPL7157137468	109543012	109543381	Pool 2	CCCGACCAGGTTGAGTACTTTG	GGAAAATGTTGAATTTGGCATGTTCATG
AMPL7157137470	109543316	109543658	Pool 1	TCTCTTCTGATCTATAGAGTGCCTTGT	ATCAGAACCCTCCAGGCAAGAT
AMPL7157137472	109543604	109543879	Pool 2	GACCACCTTTGCCTTCATTCAG	GAATCCCAGAAGAAAATGACCTCTACA
AMPL7157137496	109543819	109544110	Pool 1	CTGCCTTTGCATATCCTGCTTTC	CTACCTCTACCCTCAGTAGCTGTAC
AMPL7157137497	109544055	109544426	Pool 2	AGCAGCTTAGAGTTGCAGCA	CGGCCAGAATATTATAATCTTATGGGAC
AMPL7157137498	109544759	109544904	Pool 1	TCTTTTCTATGTTTAGATACACAAATGCCTACC	TGGGATAACCTACCACTACACACC
AMPL7157137499	109544845	109545139	Pool 2	GTAGACCAGGAGCAATAAGCTACAC	CTGCTACTCAGGTGACTAAGGTG
AMPL7157137500	109544978	109545328	Pool 1	GCATCCCTGTCATTAGGTGATGT	AAAAAGAGGAATAGTCAGTGGCC
AMPL7157137501	109545300	109545658	Pool 2	ACACCTGGCCACTGACTATTC	AAAAATCTCGCCACAAAAATGGAAT
AMPL7157137502	109545625	109545750	Pool 1	CGGCCAACTATTCCATTTTTGTG	GCCATTTACTTTCCATTTTGTGCCTTT
AMPL7157137503	109545683	109546057	Pool 2	TCTTCCTTCTTAGGAGCTGTAAGACTATTC	TCATAAGCCAGGGACCAGTTGTA
AMPL7157137504	109545997	109546370	Pool 1	AGTGTGTCAGTCCCTTTATATGTACCT	CCACCTCCTAGGCTCAAGTTACT
AMPL7157135550	109546551	109546823	Pool 2	TCAGTCTTCTGTGTAGCTAGGACTG	GGGCAAAACAAAGCTTCTTTTACAATGA
AMPL7157135551	109546757	109547023	Pool 1	AGTACAGTGTTCAATGAGTTACATGAGG	AGGACACCATGATGCATACACAC
AMPL7157137505	109546963	109547283	Pool 2	CCCATCCATTGTGTTAATAAGTTGAGG	AAAAGGAAGGACTGTGGCTGTT
AMPL7157137485	109547236	109547610	Pool 1	GGGAATCGCACAGGGTCTTA	GTGGCAGATTTGGGACTTTTTAAATTTTT
AMPL7157137486	109547582	109547815	Pool 2	AAAAATTTAAAAAGTCCCAAATCTGCCA	CGTTAACAGCAGCTTCTCAAAGG
AMPL7157119889	109547759	109548078	Pool 1	CTGGAAGGAGCTGTGATCATCAG	CCCTGTTTCACCATAAAGGCAAG
AMPL7157119890	109548024	109548389	Pool 2	GGTCAAATACTCAGTTGGCTCTCT	TCTCGGTCTCTTATCTGATAAAGCTCA
AMPL7157119865	109548325	109548515	Pool 1	GCTCAATTTCCTGATTGTAGTAGAGGT	GGACATAACCCTAGGAGAAAATCACC
AMPL7160276503	109548452	109548821	Pool 2	GGAAGAGGAGAAAAGGGAATTTTGTCT	CACAAAGGACACTTGCACAAAAC

NEIL1:chr15:75,637,831:75,647,592

Amplicon ID	Start	End	Primer Pool	Forward Primer	Reverse Primer
AMPL7160266878	75637650	75638024	Pool 1	GATGCAATTCAAAGTACAGAGCAGTG	AATATTCAGGGATACTTCCTAGGTGCT
AMPL7160266879	75637985	75638290	Pool 2	CGAAGTACAGAGCAGCACCTA	GCGCCTGGCTGGATTATATCTTTTTAATAA
AMPL7160266880	75638228	75638521	Pool 1	GGAGAGGGAAGGTGGTTTCATC	CAGGCTACAATCATGGGTCATTG
AMPL7160088421	75638380	75638637	Pool 2	TAGTGAGACCACATCCTTACAAAACATG	GCCCAGCCAACAGTAACTTCTTAA
AMPL7158763775	75639030	75639293	Pool 1	AGACAAATCATGACATTTCATACCCAATACT	CTGGAGTATAGCTGCGGGTTAG
AMPL7157141564	75639238	75639567	Pool 2	CCTGGGTTTGCTTCTCTGTCTTC	TTAAGCACCAACGAACAGTCAGA
AMPL7157141565	75639513	75639803	Pool 1	CAGGCCGTGATGATGTTTGTT	GTGGCGGAAGAACCTGACTT
AMPL7157141567	75639776	75640069	Pool 2	GCAGTGGGAAGTCAGGTTCTTC	GCTCATCTTCTCCTTAGCCCTACA
AMPL7155948367	75640017	75640384	Pool 1	CGTGCAGGATCGGGTTGT	GGGTATCGACTCCTCACCTC
AMPL7157141594	75640649	75640928	Pool 2	GGCCGAGCTTCGCTCTTT	CGGCTGCAGCCTTACTTCTT
AMPL7157141595	75640876	75641194	Pool 1	CGGAGATGAGATCCGTGAGTGA	CGGTGTGGACTTTGTGAGGTAG
AMPL7157141569	75641137	75641346	Pool 2	CTCCGAGTTCTCCTCTAAAAATGGG	GGCTGACAGAGGACTTCTCCA
AMPL7155947801	75641275	75641600	Pool 1	CCAGCCAGTTTGTGAATGAGG	GCGGATGTCCACGAAACATAGG
AMPL7157141596	75641540	75641873	Pool 2	CCTGCGCTTTTACACGGC	CCCAATTAGTTTTGAGTTTCCATGC
AMPL7157141571	75641842	75642215	Pool 1	CCAGTGGGCATGGAAACTCAA	CCCATCCTCCTCTTGGCTCT
AMPL7157136034	75642198	75642401	Pool 2	AGCCAAGAGGAGGATGGGTAG	GAAGCCATGTCATGAGATAGTCCC
AMPL7157141573	75642345	75642686	Pool 1	GGCTGAGCGGAATTGTTTCAAG	CCTATCCTATGTGAACATCAGTTCTGTG
AMPL7157141574	75642622	75642981	Pool 2	GTACAGCAGTGAACAAGATGAATGTG	TAAGGCACCAGAGACATCTCCAA
AMPL7157141575	75642924	75643270	Pool 1	TGACCCTTAGGAGACTTGGATACC	GTGATGAGTGGGTCAGAAGACA
AMPL7157141576	75643216	75643491	Pool 2	CCAGGCTACCCTTGTACCTTCA	TGGGTTAGAGACCATATGTGGGT
AMPL7157141577	75643434	75643808	Pool 1	AGGATCTTGAGAAGAATGGACCCT	CATCTCTACCAAAAATTTAAGAAGTAGCCAG
AMPL7157141578	75643733	75644083	Pool 2	CAACCTCCTGATTAACTGGAACCA	GCCTGAAGCAGCAGAATCTCTG
AMPL7157141598	75643903	75644277	Pool 1	CACCAGCCTCAAGTTACTTTAGATCA	TCATTGTGAAAGGACGGCCA
AMPL7157141581	75644256	75644578	Pool 2	CCTGGCCGTCCTTTCACAA	TGCTGACCGGTACAGGATCT
AMPL7157141582	75644524	75644898	Pool 1	GTTCTTCAATGGCATTGGCAACTA	GTAAAGGCAGCTGACCTTCTCT
AMPL7157141583	75644845	75645205	Pool 2	GGATGAACTGCCCAAAGTCTGA	CCCAGCCTGGAAACACTATTGA
AMPL7157141584	75645157	75645528	Pool 1	ACTCAATGGACTAGGCCTCCT	TGGGCAGCTTGGAGGAAAC
AMPL7157134531	75645501	75645659	Pool 2	AAAAACAGTGTTTCCTCCAAGCT	CCCAGTTTGAGGCTAGTCCAG
AMPL7157141586	75645585	75645939	Pool 1	CACACACGCACGTTTATATATATTTTTGGT	TGTGACCTCAGCTGGTTTCTTC
AMPL7157141587	75645880	75646238	Pool 2	TAATCCCACTCCAGGATGGGAA	GTCTCTGCCTGTGTGAACAGTAG
AMPL7157141588	75646185	75646483	Pool 1	ATGGCCGTACCATCTGGTTC	CCCTTAGGAAGCCTCTAGGACA

NEIL1:chr15:75,637,831:75,647,592

AMPL7155947798	75646429	75646800	Pool 2	CAGCTTTGCTGATGTGGACTCT	CTTGGATTTCTTTTTGCGGGACTT
AMPL7157141589	75646735	75647070	Pool 1	GATAGGACCCTCCAACTCCAAC	TGGGTTGCAGTCCTCTTAGGAA
AMPL7157136027	75647019	75647250	Pool 2	CTTCCAGGACACGAAGGGCAAA	GCCTTTCTCCCACCTTTGTACAC
AMPL7153478917	75647190	75647464	Pool 1	GCTCTACAGGAGAGGATGGGAT	TCAGATATTGCCTGCTCCCAAAAA
AMPL7157141590	75647406	75647718	Pool 2	TTCTTATTGTCTTGCCCTGCATCT	GATAGAGTCAGAGCCTTGAGGACT

APOE:chr19: 45,409,039:45,412,650

Amplicon ID	Start	End	Primer Pool	Forward Primer	Reverse Primer
AMPL7152996218	45409014	45409161	Pool 1	GGGAGCCCTATAATTGGACAAGT	AGGGTCCCAGCTCTTTCTAGAG
AMPL7154479522	45409755	45410121	Pool 2	CAACAAGGCTTGGAAGGCTAAC	CAAAACAACTTCTAACTCCTATCTCAAGGA
AMPL7154479521	45410987	45411342	Pool 1	CCTGACCCACCTTGAACTTGTT	ATGGGAAGAGGAAGCTAGAACCA
AMPL7154479523	45411591	45411842	Pool 2	CCTCCTAGCTCCTTCTTCGTCT	CCAGTTCCGATTTGTAGGCCTT
AMPL7154479524	45411792	45412076	Pool 1	CGCTGATGGACGAGACCA	TCTGCAGGTCATCGGCATC
AMPL7154443516	45412028	45412367	Pool 2	CTGCGCAAGCTGCGTAAG	CCTGCAGGCGTATCTGCTG
AMPL7155127581	45412292	45412429	Pool 1	CTGGACGAGGTGAAGGAGCAG	GCGCTGCATGTCTTCCAC
AMPL7155127582	45412385	45412685	Pool 2	CGCCTCAAGAGCTGGTTCG	GGGCTTAGAGGAAATCACAGGG

APOE PROMOTER:chr19:45,408,011:45,409,011

Amplicon ID	Start	End	Primer Pool	Forward Primer	Reverse Primer
AMPL7160266886	45407922	45408265	Pool 1	GGCTGTCCTCCATGCTCAATAC	GCAAATGGGACCTGGGTTTTAATCA
AMPL7160266817	45408209	45408387	Pool 2	TGTGCTCAAGGTCACAACCAA	CCTCAGCAAGAGGGAGACTGT
AMPL7157901648	45408366	45408740	Pool 1	GACAGTCTCCCTCTTGCTGAG	CCTGGATGAATGTAATCTGGAGAGG
AMPL7160266885	45408668	45409040	Pool 2	CCTAGCCCTACTTTCTTTCTGGGAT	CCAGACTTGTCCAATTATAGGGCT

8. CURRICULUM VITAE

Personal Information

Name	Tuğçe	**Surname**	Ertüzün
Place of Birth	Altındağ, Ankara	**Date of Birth**	18.12.1992
Nationality	TC	**Telephone number**	+905462404240
E-mail	tuce.ilhan@gmail.com		

Education

Level	Institution Name	Graduation Year
Doctor of Philosophy	Acıbadem Mehmet Ali Aydınlar University	2020
Master of Science	Acıbadem Mehmet Ali Aydınlar University	2016
Undergraduate	Istanbul Bilgi University	2014
High School	Büyük High School	2012

Work Experience

Position	Corporation	Duration
Protein Analysis Specialist	Turgut Ilaclari A.S.	08.2019-
Research Assistant	Acıbadem Mehmet Ali Aydınlar University	05.2015-07.2019
Laboratory Intern	ETH Zurich - Department of Biosystems Science and Engineering	06.2013-08.2013

Foreign Languages

Language	Reading*	Speaking*	Writing*
English	Advanced	Advanced	Advanced
German	Beginner	Beginner	Beginner
French	Beginner	Beginner	Beginner

***** Evaluated as advanced, good, intermediate, beginner

Foreign Language Exam Results [#]

YDS	YÖK-DİL	IELTS	TOEFL IBT	TOEFL PBT	TOEFL CBT	FCE	CAE	CPE
85,000	90							

[#] All successful exams should be enrolled.

[#] YDS: Yabancı Dil Sınavı; YÖK-Dil: Yükseköğretim Kurumları Yabancı Dil Sınavı; IELTS: International English Language Testing System; TOEFL IBT: Test of English as a Foreign Language-Internet-Based Test TOEFL PBT: Test of English as a Foreign

Language-Paper-Based Test; TOEFL CBT: Test of English as a Foreign Language-Computer-Based Test; FCE: First Certificate in English; CAE: Certificate in Advanced English; CPE: Certificate of Proficiency in English

Other Exams

Name of the Exam	Quantitative	Equally Weighted	Verbal
ALES**	78,44123	79,61634	70,84783

**ALES: Akademik Personel ve Lisansüstü Eğitimi Giriş Sınavı

Computer Skills

Program	Ability to Use
Microsoft Office	Advanced
MATLAB	Beginner
GraphPad Prism 6	Beginner
CLC Genomics Workbench	Intermediate

* Evaluated as advanced, good, intermediate, beginner

Projects

Name of the Project	Institution	Position	Years
Sporadik Alzheimer Hastalığı için Yeni Bir Genetik Faktörü Olarak Baz Eksizyon Tamir Genlerinin Genetik Varyantlarının Taranması ve Değerlendirilmesi TUBITAK 1001, Project No: 114Z875	Acıbadem Mehmet Ali Aydınlar University	Research Scholar	05.2015-05.2018
Development of the original 2-indolinone compounds as anti-interleukin 1 and chemotherapeutic drugs Sub-project 2: DNA repair inhibitors to be developed as a new drug for the treatment of hereditary nonpolyposis colorectal cancer (MLH1 gene defect). TUBITAK 1003, Project No: 215S614	Acıbadem Mehmet Ali Aydınlar University	Research Scholar	01.2016-01.2019
Determination of Small Molecule Inhibitors Blocking Mitochondrial Base Repairing Mechanism and Their Activities by Targeting DNA Polymerase Gamma Enzyme for Cancer Therapy. TUBITAK 1001, Project No: 212T026	Acıbadem Mehmet Ali Aydınlar University	Research Scholar	05.2015-11.2016

Scholarships

Type	Institution	Years

Full scholarship	Acıbadem Mehmet Ali Aydınlar University, Medical Biotechnology PhD Program	2016-2020
Full scholarship	Istanbul Bilgi University, Genetics and Bioengineering	2010/2014
Full scholarship	Büyük High School	2006-2010

Conferences

Name of Conference	Position	Institution/Place	Duration/ Year
6th International Congress of the Molecular Biology Association of turkey	Participant	İzmir	5-8.09.2018
Genome Maintenance, DNA Repair and Cancer	Organization committee and participant	Acıbadem Mehmet Ali Aydınlar University	26-28.04.2018
EMBO Young Scientists' Forum	Participant	Boğaziçi University	05-06.01.2018

International and National Courses and Certificates

Name of Certificate	Institution	Year
Certificate of Experimental Animal Use	Acıbadem Mehmet Ali Aydınlar University, DEHAM	2018

PUBLICATIONS:

Ertuzun T, Semerci A, Cakir ME, Ekmekcioglu A, Gok MO, Soltys DT, Sezerman U, Muftuoglu M. Investigation of base excision repair gene variants in late-onset Alzheimer's disease. PLoS One. 2019;14(8):e0221362. doi: 10.1371/journal.pone.0221362. PubMed PMID: 31415677; PubMed Central PMCID: PMCPMC6695184.

PROCEEDINGS

1. Selcuk Keskin, Berna Somuncu, Merve Antmen, Yesim Saglican, **Tugce Ertuzun**, Mustafa Bilal Tuna, Can Obek, Umit Ince, Ali Riza Kural, Meltem Muftuoglu. Kasa invaza olmayan mesane tümöründe artmış apurinic/apirimidinik endonükleaz aktivitesi. 4.Üroloji Cerrahi Kongresi, 31, 2018.
2. Keskin S, Somuncu B, Antmen M, Saglican Y, **Ertuzun T**, Tuna MB, Obek C, Ince U, Kural AR, Muftuoglu M. Enhanced Apurinic/Apyrimidinic Endonuclease Activity in Non-Muscle Invasive Bladder Cancer. 36th world congress of Endourology, 6, 2018.
3. İsmail Öğülür, **Tuğçe Ertüzün**, Ahmet Özen, Emel Uyar, Ayça Kıykım, Dilek Başer, Gözde Yeşil, Hacer Aktürk, Ayper Somer, Meltem Müftüoğlu, Elif Karakoç Aydıner, Safa Barış. Ataksi telenjiektazi hastalarının ebeveynlerinde enfeksiyonlar ve lenfosit alt grup bozuklukları. XXV.Ulusal Alerji ve Klinik İmmünoloji Kongresi. S-0-11, 2018.
4. Ogulur I, **Ertuzun T**, Ozen A, Uyar E, Kiykim A, Baser D, Yesil G, Akturk H, Somer A, Muftuoglu M, Karakoc-Aydiner E, Baris S. Infections and Lymphocite

Subset Alterations in Ataxia Telangiectasia Mutated Kinase Carriers. 18th Biennial Meeting of the European Society for Immunodeficiencies, ESID8-0517, 2018.

TUĞÇE ERTÜZÜN

T.C. ACIBADEM MEHMET ALİ
AYDINLAR UNIVERSITY
INSTITUTE OF HEALTH SCIENCES

DOCTORATE
THESIS

ISTANBUL 2020