REPUBLIC OF TURKEY

ACIBADEM MEHMET ALİ AYDINLAR UNIVERSITY

INSTITUTE OF HEALTH SCIENCES

INVESTIGATION OF *UNG* GENETIC VARIANTS IN SPORADIC ALZHEIMER DISEASE

Aslı SEMERCİ

MASTER THESIS

DEPARTMENT of MEDICAL BIOTECHNOLOGY

SUPERVISOR

Prof. Dr. Meltem MÜFTÜOĞLU

İSTANBUL-2018

REPUBLIC OF TURKEY

ACIBADEM MEHMET ALİ AYDINLAR UNIVERSITY

INSTITUTE OF HEALTH SCIENCES

INVESTIGATION OF *UNG* GENETIC VARIANTS IN SPORADIC ALZHEIMER DISEASE

Aslı SEMERCİ

MASTER THESIS

DEPARTMENT of MEDICAL BIOTECHNOLOGY

SUPERVISOR

Prof. Dr. Meltem MÜFTÜOĞLU

İSTANBUL-2018

DECLARATION

I hereby declare that, this thesis has been written by me based on the data obtained in line with the scientific rules and ethical principles of responsible conduct of research. All information, data, comments, analyses have been collected and processed through scientific, academic writing style, and literature used have been duly shown by giving reference to the original sources in accordance with the publication ethics. I also announce and emphasize that I have not violated any rules secured by patent and copyrights whilst the conduct and writing of this research.

29.06.2018

ASLI SEMERCİ

ACKNOWLEDGEMENTS

I would like to thank my supervisor Prof. Dr. Meltem Müftüoğlu for her continuous guidance and support during my master studies. It is an honor and great pleasure for me to work with her.

This study was supported by the Scientific and Technological Research Council of Turkey (TUBITAK), grant no: 114Z875. I would like to thank TUBITAK for its financial support.

I would you like to Prof. Dr. Uğur Sezerman for his help and support on the bioinformatics analysis. Also, I would like to thank Prof. Dr. Batu Erman, and Medeniyet University Research and Training Hospital Neurology Department doctors and nurses for their support. I owe special thanks to my colleague Tuğçe Ertüzün and Acibadem Mehmet Ali Aydınlar University research laboratory members for their help during my master studies.

Of course, I am grateful to my friends who have become my life during my master education. Life has been better with them.

Last but definitely not the least I am grateful to my family for the endless support they have given. They have always been with me, and it'll not change by anytime.

Aslı Semerci

LIST OF CONTENTS

LIST OF SYMBOLS/ABBREVIATIONS

AD: Alzheimer's disease

APE-1: AP-endonuclease

APOE: Apolipoprotein-E

APP: Amyloid-precursor protein

Aβ: Amyloid-Beta

BER: Base excision repair

CI: Confidence interval

CLU: Clusterin

CR1: Complement C3b/C4b Receptor 1

CSF: Cerebrospinal fluid

DNA: Deoxyribonucleic acid

dsDNA: Double strand DNA

EDTA: Ethylenediaminetetraacetic acid

ETS: Electron transport system

EXOG: Exo/Endonuclease G

Fapy-A: Fapy-adenine

Fapy-G: Fapy-guanine

FDR: False-discovery rate

FRMD4A: FERM Domain-Containing Protein 4A

GWAS: Genome-wide association study

Ig: Immunoglobulin

IGV: Integrative Genomics Viewer

Ion-PGM: Ion Torrent Personal Sequencing Machine

LIG-1: DNA Ligase-1

LIG-3: DNA Ligase 3

MAF: Minor allele frequency

MCI: Mild-cognitive impairment

MMR: Mis-match repair

MMSE: Mini mental state examination

mtDNA: Mitochondrial DNA

nDNA: Nuclear DNA

NEIL1: Nei-like DNA glycosylase 1

NER: Nucleotide-excision repair

NGS: Next generation sequencing

OD: Odds ratio

OGG-1: 8-dihydroguanine DNA glycosylase-1

PARP-1: Poly (ADP-Ribose) polymerase 1

PCR: Polymerase chain reaction

PET: Positron electron transmission

POL γ: DNA polymerase gamma

POLβ: DNA polymerase beta

PSEN: Presenilin

RCF: Relative centrifugal force

RNase A: Ribonuclease A

ROS: Reactive oxygen species

SMUG-1: Selective monofunctional uracil DNA glycosylase-1

SNP: Single nucleotide polymorphisms

SNV: Single nucleotide variation

ssDNA: Single strand DNA

TDG: Thymine DNA-glycosylase

UNG: Uracil DNA glycosylase

UTR-3: Three prime untranslated region

UTR-5: Five prime untranslated region

UV: Ultraviolet

XRCC1: X-Ray repair cross-complementing 1

LIST OF FIGURES

LIST OF TABLES

SUMMARY

The oxidative DNA damage accumulation and defect of its repair through "base excision repair (BER)" mechanism is crucial for the sporadic AD's etiology and pathogenesis. Uracil DNA glycosylase (UDG) incises uracil lesion from DNA involved in the first step of both nuclear and mitochondrial BER. The biochemical, cellular, molecular and behavioral studies are done with the post-mortem brain tissues of sporadic Alzheimer patients, Alzheimer mouse models and cell lines revealed the strong relation between BER deficiency due to the decrease in the expression levels and activities of UDG, and sporadic AD pathogenesis. In addition, UNG gene silencing makes neurons susceptible to amyloid AβPP toxicity. Therefore, in this study, *UNG* gene was selected as potential genetic factor for sporadic AD. Thus, *UNG* gene covering promoter, exon and intron regions in sporadic Alzheimer patients and healthy controls were sequenced using Ion-PGM NGS technology and the significant *UNG* gene variants were confirmed with Sanger sequencing. This study revealed two new genetic risk factors for sporadic AD ($p<0.05$). Furthermore, this study also demonstrated that the decrease in the expression, protein and activity levels of UDG in sporadic AD is not due to *UNG* gene variant(s).

Key Words: Base excision repair, Genetic variant, Next generation sequencing, Sporadic Alzheimer disease, Uracil DNA glycosylase.

ÖZET

Sporadik Alzheimer Hastalığında UNG Genetik Varyantlarının Araştırılması

Oksidatif DNA hasarının birikmesi ve bu hasarı onaran baz eksizyon tamir (BER) mekanizmasındaki bozukluk sporadik Alzheimer hastalığının (AH) etiyolojisi ve patogenezinde rol oynamaktadır. Urasil DNA glikozilaz (UDG) urasil DNA lezyonunu keser ve nükleer ve mitokondriyal BER yolağının ilk basamağında yer alır. Sporadik Alzheimer hastalarının post-mortem beyin dokularında, Alzheimer fare modellerinde ve hücre hatlarında yapılan biyokimyasal, moleküler, hücresel ve davranış çalışmaları, sporadik AH'nın patogenezi ile UDG'nin ekspresyon ve aktivitesindeki azalmanın neden olduğu BER hasarı arasındaki güçlü ilişkiyi ortaya çıkarmıştır. *UNG* geninin sessizleştirilmesi sinir hücrelerini amiloid AβPP toksisitesine duyarlı kılmaktadır. Bu nedenle, bu çalışmada, *UNG* geni sporadik AH için genetik risk faktörü adayı olarak seçilmiştir. Sporadik Alzheimer hasta ve sağlıklı kontrol gruplarında *UNG* geni (promoter, ekzon ve intron bölgeleri) Ion-PGM NGS teknolojisi kullanılarak dizilendi ve önemli *UNG* gen varyantları Sanger sekanslama ile doğrulandı. Bu çalışma, sporadik AH için iki yeni genetik risk faktörü olduğunu ortaya çıkarmıştır ($p < 0.05$). Ayrıca, bu çalışma aynı zamanda, sporadik AH'de UDG'nin ekspresyon, protein ve aktivite düzeylerindeki azalmanın, *UNG* gen varyant (lar) ından kaynaklanmadığını da göstermiştir.

Anahtar Sözcükler: Baz eksizyon tamiri, Genetik varyant, Sporadik Alzheimer hastalığı, Urasil DNA Glikozilaz, Yeni nesil dizileme.

1. BACKGROUND OF THE STUDY

Aging and environmental factors are crucial in the sporadic AD's etiology (1). Since it has been shown that early diagnosis slow the progression of sporadic AD and reduce its symptoms, screening studies for the identification of genetic risk factors for sporadic AD has increased. Nerve cells are under constant attack of oxidative damage, because of their high energy requirements. The oxidative DNA damage accumulation and the defect in BER mechanism for repairing oxidative DNA damage especially in nerve cells are crucial for the etiology and pathogenesis of various neurodegenerative diseases, including sporadic AD (2,3). Thus, in recent years, BER genes have become the focus of research to determine genetic risk factors for sporadic AD, but BER gene/gene variant(s) related to sporadic AD has not been determined yet.

UDG is a monofunctional DNA glycosylase involved in the first step of both nuclear and mitochondrial BER. *UNG* encodes both nuclear and mitochondrial UDG isoforms formed by alternative splicing (4,5). The lack of UDG protein due to *UNG* gene silencing in rat hippocampal neurons caused neuronal death by inducing neuronal apoptosis, suggesting that this protein plays a crucial role in the neuronal development (6). It has been shown that folic acid deficiency and high homocysteine levels (7), which increase the risk of Alzheimer's disease, enhance DNA damage in nerve cells, in particular uracil misincorporation of DNA and hypomethylation of DNA, and thus make neurons susceptible to amyloid AβPP toxicity (8). The studies in Alzheimer's mouse models and cell lines have shown that the damage to the BER pathway due to the decrease in uracil incision activity of the UDG protein also contributes to the increase in neuronal hippocampal neurodegeneration and neuronal AβPP toxicity (6,8). The studies in post-mortem brain tissues of sporadic Alzheimer's patients have shown that UDG expression levels are significantly reduced and also the recognition of uracil lesion and incision activities of UDG decreases (9). In these studies, the activity and protein levels of UDG were low both in inferior parietal lobule ("the most affected region") and in the cerebellum (less affected region). The presence of BER lesions due to deficiency in UDG activity in

post-mortem brain tissue of patients diagnosed with mild cognitive impairment, which is considered to be a transition period between normal aging and dementia, indicates that defects in BER mechanism can also be observed in the early onset of sporadic AD. This suggests that BER pathogenesis is not only for the neuropathological areas, it is also general character of Alzheimer brain regions (9–13).

Therefore, in this thesis study, *UNG* gene is selected as a genetic risk factor candidate for the sporadic AD. It is also investigated whether the significant decrease in the protein level, expression and activity of UDG in sporadic AD patients is due to genetic variant(s) in *UNG* gene. Thus, *UNG* gene (promoter, exon and intron regions) in sporadic AD patients and healthy controls were sequenced using Ion-PGM NGS technology and the significant *UNG* gene variants were confirmed with Sanger sequencing.

2. INTRODUCTION

2.1.Alzheimer Disease

AD is a complex progressive neurodegenerative disorder. The main cause of AD is the accumulation of incorrectly folded amyloid proteins for example amyloid beta (Aβ) in the extracellular area and hyperphosphorylated tau proteins in the cytoplasmic area of the brain. The amyloid precursor protein (APP) is produced in the brain during the normal physiological conditions and cleaved by proteases to generate functional amyloid proteins. The wrong cleavage and misfolding of the amyloid proteins (Aβ40 and Aβ42) cause a neurotoxic effect in the brain. Plaque formation of Aβ disconnects the nerve cells transmission. The degradation of hyperphosphorylated tau is inhibited by Aβ. Thus, the tau proteins accumulate in the cytoplasm of nerve cells and cause a neuronal damage. Early signs of AD are observed in the frontal and temporal lobes of the brain. Brain deformities diffuse to other brain regions such as the hippocampus. As the disease progresses, more and more areas of the brain become damaged (Figure 2.1) (1).

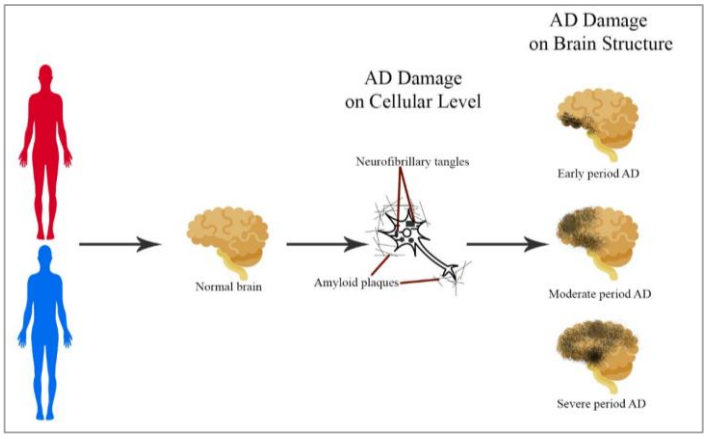


Figure 2.1: The schema for AD progression in the patient's brain. Nerve cells' dis-functionality is caused by Aβ plaques and hyperphosphorylated tau proteins. Different stages of the disorder are affected to different brain regions (1).

"Familial AD" and "sporadic AD" are major AD types. The familial AD is a Mendelian hereditary genetic disorder. Early onset familial AD (age at onset < 65

years) is a rare form of AD, and accounts for approximately 5% of all cases of AD (1). Three genes cause familial AD: *"APP*, presenilin-1 (*PSEN-1*), and 2 (*PSEN-2*)" (14). The sporadic AD is the most common type of the disease that found in ~95% of total AD population. Aging is the most common risk factor for late onset sporadic AD and most cases begin usually above the age of 65 years (15). Both sporadic AD and familial AD have same pathology (16). There are different hypothesis for the mechanisms of AD such as cholinergic hypothesis, amyloid cascade (ACH) hypothesis, presenilin hypothesis and APP matrix approach. The cholinergic hypothesis is the first hypothesis for explaining the AD mechanism. In this hypothesis, defective mechanism of acetylcholine production leads to Aβ aggregation in the "central nervous system (CNS)". According to this hypothesis, commercial drugs for example L-dopamine (donepezil) are designed and started to use for the treatment of AD (17). ACH hypothesis is based on the production of mutant forms of APP and its deposition in the CNS. According to this hypothesis, due to increase in Aβ40-42 levels and the failure of Aβ clearance, Aβ plaques accumulate in the brain of AD patients (16). The Aβ accumulation causes neuronal and synaptic dysfunction, nerve cells loss and dementia in AD patient. ACH is the most dominant disease progression model. *APP* has been expressed at high levels in the CNS of AD patients. Turnover mechanism of APP proteins disrupted and incorrect protein production is induced by secretases and presenilins. This circumstances lead to the phosphorylation of tau protein and production of tangle formation (18,19). Presenilin hypothesis is related to familial AD. Because of mutations in *PSEN* gene, PSEN protein function decreases. The alteration in signaling pathways cause gliosis, tau hyperphosphorylation and neurodegeneration in the brain. "APP matrix approach" is based on the dynamic complexity of APP proteolytic system. Full length APP is cleaved by the cellular matrix proteins. Any mutation in the *APP* gene causes major regulatory changes in the protein cleavage and then dementia progression started. The failure of this system deregulates the cholesterol homeostasis, immune signaling and synaptic plasticity (16). In addition to described hypothesis there are tau and tangle hypothesis. This hypothesis provide an information about relationship between microtubule stabilization and tau hyperphosphorylation, the wnt signaling pathway pathology that inhibits the

“glycogen synthase kinase-3” and cause tau polymerization (18), and the multiple variations (20) have been described for understanding of AD mechanism.

There are different cognitive impairment tests for the diagnosis of AD: A. Investigation of Aβ42 protein accumulation in the brain by “positron electron transmission analysis (PET scanning)”, and protein analysis in the cerebrospinal fluid (CSF). B. Investigation of hyperphosphorylated tau protein accumulation performed by PET scanning and protein analysis in the CSF. C. Investigation of nerve cells damage performed by labeled-PET scanning, “magnetic resonance imaging (MRI)”, and protein analysis in the CSF (21). In addition to these tests, psychological tests are performed for diagnosis of the familial and sporadic AD. The commonly used psychological test is the Folstein test (also known as the “mini mental state test, MMSE”) and it has variable questions that change with the cultural differences. It is an important disorder classification criterion for estimation of the patients’ physiological conditions. This test separates the patients into five groups: “healthy, suspicious-AD or MCI, mild-AD, moderate-AD, and severe-AD” (22).

2.2. DNA Repair Pathways

The genome is damaged by “endogenous and exogenous agents” and thus generates genomic instability. DNA damaging agents generate “reactive-oxygen species (ROS) such as hydroxyl radicals, superoxide radicals or hydrogen peroxide” (Figure 2.2) (23,24). Cells have specialized “DNA repair pathways” to protect the genome from the harmful effects of DNA damaging agents. Five major DNA repair pathways: “base excision repair (BER), nucleotide excision repair (BER), mismatch repair (MMR), homologous recombination (HR) and non-homologous end joining (NHEJ)”. DNA damage can be generated by environmental agents, for example radiation, diet or accumulation of metal ions (25) or endogenous agents that are byproducts of normal metabolism, for example, mitochondrial electron transport chain. If DNA damages are not corrected in the genome, it can cause genomic instability. There are several different DNA repair-deficient syndromes that can be caused by mutations of DNA repair genes. For example, xeroderma pigmentosum is caused by mutations of NER genes (26,27).

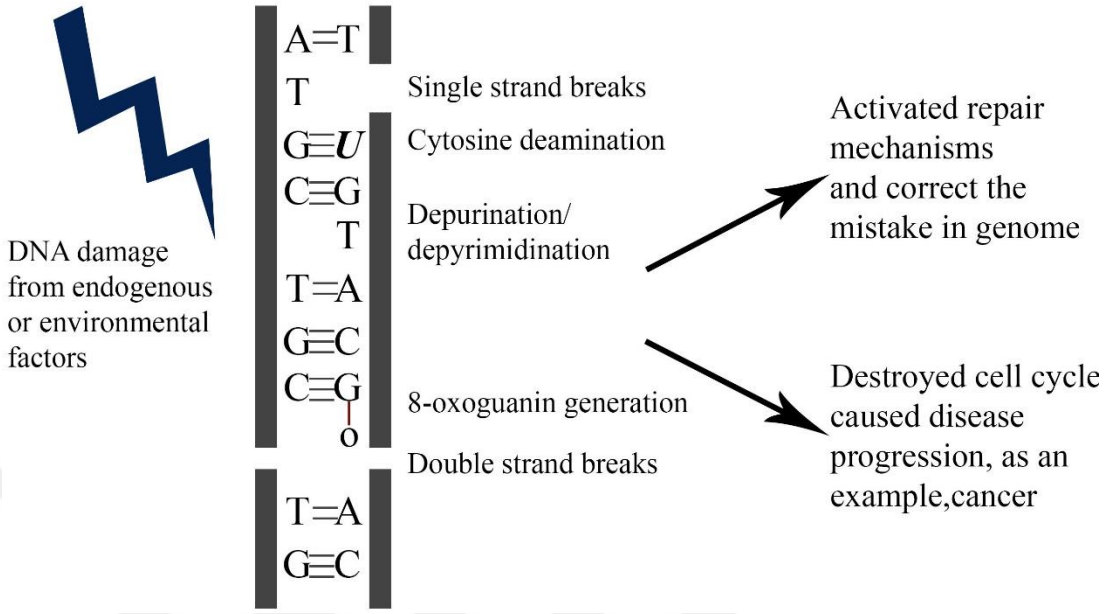


Figure 2.2: Types of DNA damages (23–26).

MMR pathway has been discovered in the *E. coli*. MMR pathway repairs mis-incorporated bases, DNA loop formations, insertions and deletions during DNA replication (28–30). Deficiency in MMR pathway causes cancers and neurodegenerative disorders (27,28). UV is one of the most important DNA damaging agent and cause cyclobutane pyrimidine dimers and photoproducts coming from UV, repaired by NER pathway. NER pathway was discovered in prokaryotes as a first repair system with Uvr proteins and this mechanism exists in nearly all types of living organisms (29,30). Dysfunctional NER pathway is associated with several types of disorders such as skin cancer, Xeroderma pigmentosum (Global genome NER) and Cockayne syndrome (Transcription Couple NER) (30–32).

2.2.1. Base excision repair pathway

The oxidized bases and single strand breaks are repaired by BER. It is started by DNA glycosylases that incise “the N-glycosidic bond between base and sugar”. DNA glycosylases are either monofunctional which has only glycosylase activity or bifunctional which has both glycosylase and AP-lyase activities (33). After the damaged base is removed by DNA glycosylases, AP endonuclease 1 (APE1) enzyme act on an abasic site (34). Then BER proceeds with either “short-patch or long-patch BER”. In “short-patch BER”, DNA polymerase β (POLβ) protein incorporates one nucleotide, and in “long-patch BER” POLβ or other polymerases adds 2-9 nucleotides. Then DNA ligases seal the nick to complete the repair (Figure 2.3) (35–38).

2.3. DNA Repair Pathways and AD

ROS accumulate into the cells with time and cause a damage on DNA, proteins, and lipids. Age-related neurodegenerative disorders are associated with decrement in nerve cells count, interactions between those cells and their functions. Also, the volume of the human brain have been decrease with age (after their 50 years old) (39). Many studies showed the decreased DNA repair capacity in the elderlies. The accumulated damages in the elderly’s genome have been associated with age-related disorders. Inefficient DNA repair causes accumulation of DNA damages especially in the un-dividing cells (like the central nervous system cells) (40).

The human brain has different types of mechanical protective mechanisms like blood brain barrier (BBB) and cranium. These systems protect the CNS from internal and mechanical damaging factors. CNS consists of generally un-divided cells, and body management is maintained by this system. Because of this, CNS cells continuously perform oxidative respiration, and the by-products arise parallel to respiration rate. Because of this, undivided cells need to have exact repair mechanisms for protection of the respiration by-products (41).

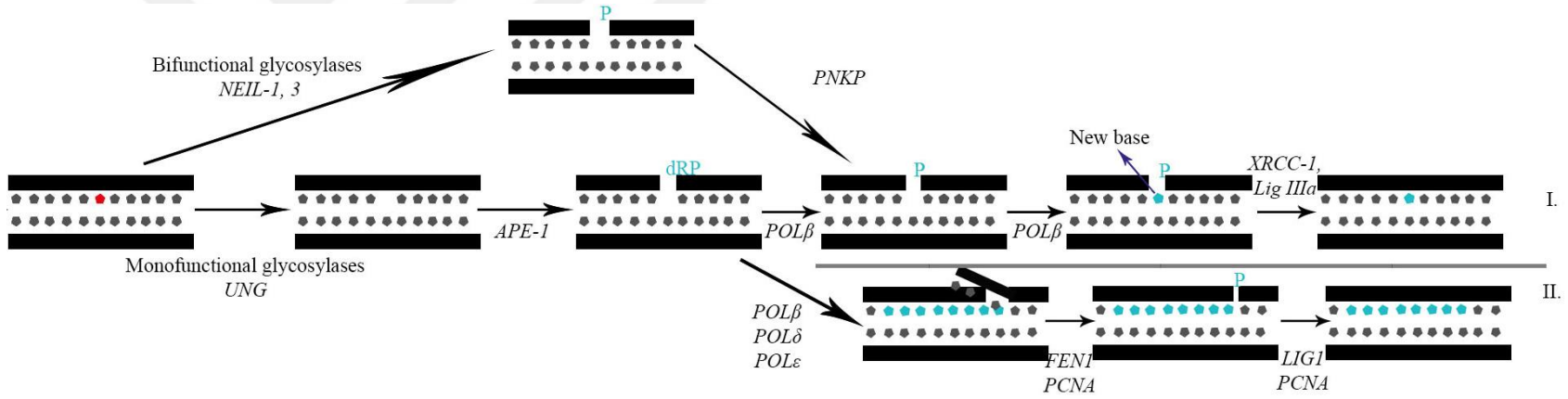


Figure 2.3: Summary of the BER pathways. The damaged base is removed with base-specific glycosylase. Two types of glycosylases are presented in the mechanism: monofunctional glycosylases are only removed the defective base. DNA-backbone opens with APE-1 activity. Bifunctional glycosylases remove the defective base, and a nick is generated in the DNA-backbone with supporting of low APE-1 activity. Opened DNA strand is repaired two different sub-pathways: "short patch (I.) and long patch (II.)". I. The genome is repaired by POLβ (with added of one nucleotide) and ligase activity. II. The genome is repaired by POLβ, POLγ, POLε (with added of 8-10 nucleotide) and ligase activity. "UNG: Uracil DNA Glycosylase, NEIL-1,3: Nei-like protein 1, 3; PNKP: Bifunctional polynucleotide phosphatase/kinase; APE-1: Human-AP Endonuclease; POL β :DNA Polymerase Beta; *POL ϱ*: DNA Polymerase-Gamma; XRCC-1: X ray repair cross complementing protein-1; LIG-1: DNA ligase-I ; LIG-3a DNA ligase-III, PCNA: Proliferating cell nuclear antigen" (33–38).

In the younger ages, metabolism by-products generate slowly, and the capacity of repair mechanisms have work high capacities. On the other hand, in the older ages, generations of metabolism by-products are generate high and so the DNA repair mechanism gets slower. Because of this, the corrupted genomic material increase in the organism (42). Also, gene expressions and protein productions of the repairing genes have also been decrease in the elderlies (43). The relationship between different types of age-related disorders and corrupted DNA regions has been investigated in the post mortem tissues (like brain). Especially, oxidized DNA bases have been increase in the MCI- and sporadic AD patients. Furthermore, changed gene expression levels are observed in the "DNA repair genes" that elderlies' and young's brain (44,45). Sporadic-AD has been associated with different types of genomic adducts that caused from ROS. Those adducts have been associated with different types of oxidized bases, SSBs, or DSBs in the nDNA or mtDNA. DNA-repair pathways are attendant to repairing of genomic damages, but developed damages leads to the accumulation of unrepaired genomic adducts in the genes of repairing pathway. So, unrepaired genomic adducts cause the production of proteins that can be over-, or fewer- secretion, nonfunctional, or pathologic. This effect is also observed in mitochondria (2).

Cellular energy production is most important function of mitochondria. Glucose interacts with oxygen, and generated energy from this interaction transferred to new phosphate bonds. Finally, adenosine triphosphate molecules have been relased from the mitochondrial system. Different types of oxidative respiration products (ROS products) have been generated and released to the mitochondrial area, and those molecules can cause genomic or proteomic defects. Also, the mtDNA doesn't have any lipid-bilayer protection, and naked DNA is utterly opened for oxidative damages. Because of this, mtDNA needs to have continuous repair in order to provide of mitochondria's and cell's integrity (11).

2.3.1. BER and Alzheimer disease

The AD studies demonstrated that the increased oxidation level in the brain of AD patients are correlated with the increased oxidized bases (46–48). Due to high

energy needs of nerve cells, oxidative respiration by-products are continuously generated inside the neuronal system (49). Also, generated DNA damage can affect cell division mechanism. Chromosomal labeling studies have been showed that the nerve cells have aneuploidic genomes in the sporadic AD patient's postmortem brain tissues (50,51). ROS can generate altered bases (oxidized or alkylated or methylated) or broken DNA strands (52,53). Lovell and colleagues have been completed serial of study for the understanding the ROS harming effects in the MCI and late-onset AD patient's brain. This study showed the effect of ROS damages to nDNA and mtDNA in the nerve cells, and these damages have been accumulate by the time in disease progression. Also, un-repaired phospholipid damages have been observed in the different regions on the postmortem brain (10,54–56). Genome methylation is another important trigger mechanism for the disease progression. Studies have been presented changes on the global methylation of DNA on the patient's genome, but any specific methylation change had not been observed in the DNA-repair pathways genes' (57,58). The uracil bases that generated from ROS attacks recognized by the glycosylases, and those adducts are removed by UDG enzyme from the DNA (41).

Respiration system's by-products oxidize DNA in the nuclear and mitochondrial genome. The researches that performed on the postmortem brain tissues showed the accumulation of oxidized bases in sporadic AD patients' brain. ROS can damage every base in the genome, and every oxidized base recognizes by the specific DNA glycosylases. Recognition of the oxidized base is initiated from the BER pathway, and the damaged base is removed from the DNA strand with the specific enzyme. DNA backbone is opened by APE1 and space formed is repaired by polymerases (POLβ in nuclear and POLγ in mitochondrial BER pathway). DNA nicks are ligated by the enzymes (*LIG1-3*) (59).

Oxidative respiration products generate different types of base adducts; "8-hydroxyl-guanine, fapy guanine, 8-dihydroxyadenine, fapy adenine, and 5-hydroxyl cytosine". The genomic material isolated from sporadic AD patients brains' and CNS's cells have been analyzed by chromatography, spectrometry and combined methods. Also, immunohistochemical analysis has been processed. All of the

findings indicated the increased of oxidized bases in the nervous system cells' genome (60,61).

Bradley-Whitman and colleagues have been completed a study for the investigation of oxidized bases in different brain regions of AD patients. The preclinical AD, sporadic AD, MCI, and age-matched control patients brain samples included in the study for analyzing the oxidized base rate in the mitochondrial and nuclear genomes. Results have been indicated that "the most adduct accumulative regions for 8-hydroxylguanine are superior and middle temporal gyri (for preclinical AD and late onset AD patients), inferior parietal lobule (for preclinical AD and MCI patients) and cerebellum (for preclinical AD, late onset AD patients and diseased controls) in the brain when compared to aged matched controls and investigated patients' brains". In addition of these results, 8-hydroxyadenine accumulates in the similar pattern in patients' brain. In spite of this, mitochondrial and nuclear genomes do not carry correlated patterns for oxidation of DNA bases (54).

BER pathway proteins are also functional in the neuronal development. The POLβ enzyme is essential for the neuronal system development. This importance is proved firstly with *in-vivo* studies. The embryonic cells carried heterozygote silenced *POLβ* gene does not repair their genome with BER pathway and their genomes' became too sensitive for alkylation (62). Besides, a mouse model is generated to investigate the importance of POLβ enzyme that combined to commonly used AD model mouse by Sykora and colleagues. So, another type of AD mouse model (*POLβ* deficient) generated for understanding the importance of BER pathway in the AD (especially sporadic AD). Generated model's genome is the carrier of high rates of DNA adducts accumulative and high predisposition for the cell death. Also, there is a new mice model that is more similar to human patients and its commonly used AD models' (59). Also, "MCI and sporadic AD patients" brains analyzed for understanding of the BER pathway's proteins level in the disease progression. UDG, OGG-1 and POLβ protein levels are decreased in the disease's progression. So, this study concluded that repair capacity is down-regulated in "sporadic AD and MCI patients" (9). DNA repair mechanisms have essential functions for neuronal signalling and production of neurotransmitters (like glutamate). Also, mitochondrial

genome damage has been induced to the production of toxic amyloid proteins'. In addition to this; mitochondrial DNA damage has been decreased with Aβ production. Also, genomic damages cause to mitophagy in the nerve cells. Mitophagy and degenerated neurotransmitter production have been observed in the DNA-repair mistakes (2).

Sporadic AD is a complex disorder. Because of this, occurred variations in the genome could generate unpredictable results. So, observed information (genomic and proteomic) should be analyzed together for generating of valid information about sporadic AD. "BER pathway's genes have been analyzed by Coppede and colleagues with the sporadic AD patient's blood samples by using methylation microarray". Results indicates that there is no significant hypo- and hyper-methylated promoter sequences in the sporadic AD samples for BER pathway's genes in compared to control samples (57). Single nucleotide variations can generate changes in the gene expressions or structure of the proteins. Findings indicate that three types of variations; un-risked variants, risk variants, and protective variants in the sporadic AD patients (Table 2.1). "Also, it has been found there is a correlation between AD progression and depression disorder". Depression disorder also frequently is shown in the elderlies. Also, strongly correlated variations have identified in the depression with the same genomic variations (63–65). Furthermore, some of the nucleotide variations have been specified to disorder. For example, in the cancer cells, functions of cellular control and repair mechanisms have degenerated, and several nucleotide variations have been associated with cancer. For example, XRCC-1 gene is carried to the rs25489 polymorphism in the different types of cancers, but this variation has not been shown in the sporadic AD patients' genomes' (66,67).

Defects in the BER pathway genes have potential role for progression of sporadic AD. Because of this, different types of studies have been continuing in order to understand the relationship between DNA-repair mechanisms and sporadic AD.

Table 2.1: The genomic alterations have been associated to sporadic AD and depression carrier patients ("Human genome, hg19, SNP144 is used for the assessment, UCSC-Genome Browser"). ↑increased risk variant, ↓protective variant (14,15).

SNP number	Gene	Variation	Function of SNP	Effects to sporadic-AD progression	References
rs4796030	*LIG-3*	A/ C	3`-UTR	No risk	(63)
rs1052133	*OGG-1*	C/G	Missense mutation Cys>Ser	↑	(63)
rs1760944	*APEX-1*	T/G	5-UTR	↑	(63)
rs4462560	*NEIL-1*	C/G	3-nearGene	↑	(63)
rs174538	*FEN-1*	G/A	5-nearGene	↑	(63)
rs1052536	*LIG-3*	C/T	3-UTR	↑	(63)
rs20579	*LIG-1*	C/T	5-UTR	↑	(63)
rs3219489	*MUTYH*	G/C	Missense mutation Gln>His	↑	(63,65)
rs1799782	*XRCC-1*	A/G	Missense mutation Arg>Trp	↑	(63,65)
rs1130409	*APEX*	T/G	Missense mutation Glu>Asp	↓	(63)
rs1136410	*PARP-1*	T/C	Missense mutation Val>Ala	↓	(63,65)
rs25487	*XRCC-1*	C/T	Missense mutation Gln>Arg	↓	(63,65)

Obtained information was classified and collected for maintaining analysis. For example, BER studies were classified as their study types: experimental observations, relations of neighbourhoods and text mining. This classified data can be compared to other data such as sporadic AD studies. This comparison demonstrates the relationship between two or more different pathways. This analysis was processed by STRING-database. Genetic network had been generated with

STRING-database for understanding the relationship between BER pathway and sporadic AD (version 10.5) (Figure 2.4).

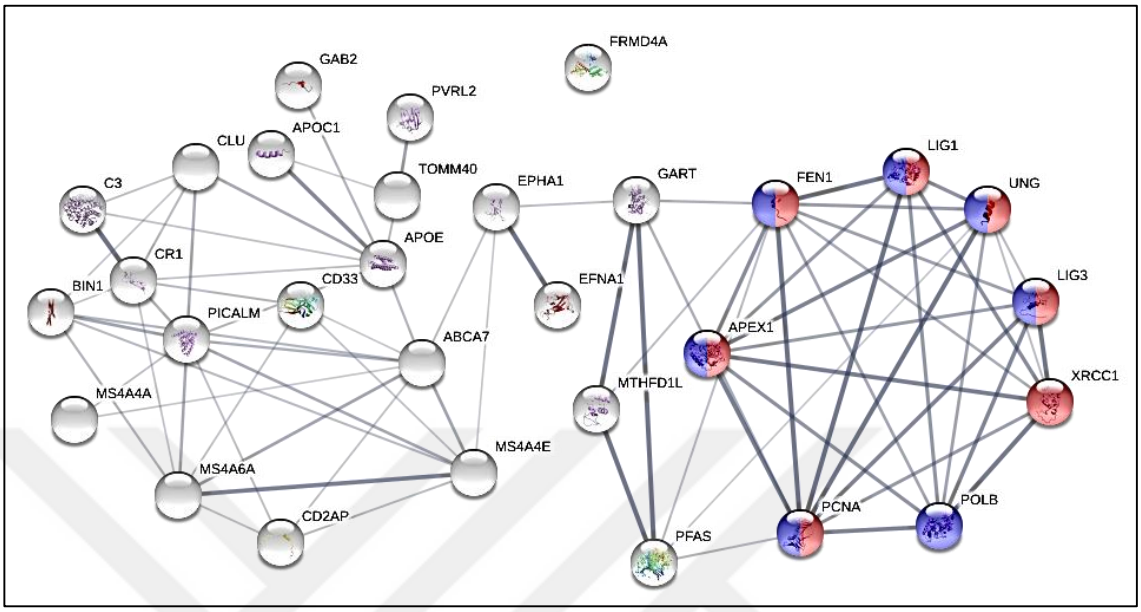


Figure 2.4: The association of BER-pathway and sporadic AD pathway related genes has been explained using with STRING-db (10.5). p-value=1.0e-16. The sporadic AD associated genes: grey, BER-pathway genes: blue and red.

2.3.1.1. Uracil-DNA Glycosylase

The oxidative deamination of cytosine to uracil generates transitions mutations (G:C>A:T) during the cell division (19). Uracil base is removed from DNA by UDG enzyme. *UNG* gene has been conserved in the evolutionary process. UDG superfamily enzymes was found in the *E. coli* and other organisms (Figure 2.5) (69–71). *UNG1* and *UNG2* genes have a GC rich and TATA fewer promoters like a housekeeping gene. Methylation regions (CpG islands) appear in the downstream of the genes (4). Mitochondrial and genomic *UNG*s have been synthesized from the same genomic region by the alternative splicing mechanism, and they are different in the N-terminal amino acid sequences (72). The *UNG2* gene is translated to 313 amino acids from six exons. The *UNG1* gene is translated to 304 amino acids from 7 exons (Figure 2.6). *UNG* gene's pseudogenes are also found in the different chromosomes (73). Uracil removal enzymes are found in different types of organisms, and these enzymes have been classified into six families according to

their sequences. Uracil removal enzymes have been found in archaea, bacterium, eukaryote and some virus types. All determined types of uracil removal enzymes have their own specific N- and C-terminal amino acid sequences, activation units and highly conserved Leu residues. All types of UDG enzymes have been used for determination and removing of uracil bases in the DNA (68,74). The UDG enzymes that determined in the *H. sapiens* and *E. coli* have been clustered into same family (family-1 of UDG).

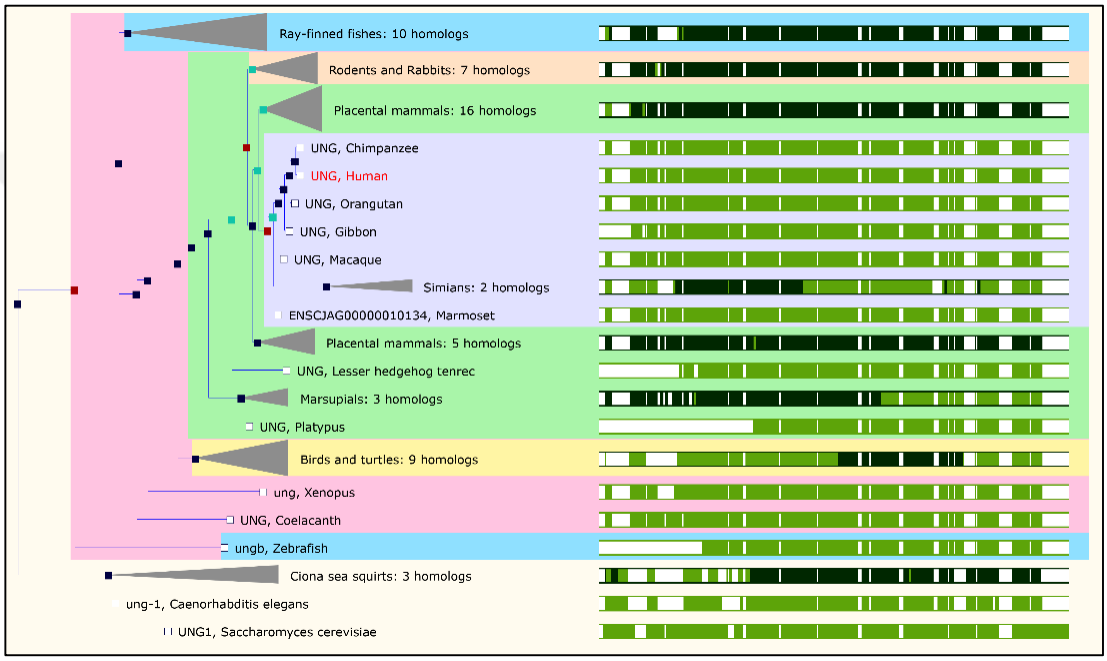


Figure 2.5: The *UNG* genes in the different organisms. The darker regions express more similar genomic sequences (>66%), the greenish regions are similar (66-33%), and the lighter regions are less similar (<33%) in the genomic sequences (hg19 is used for the comparison of the similarity).

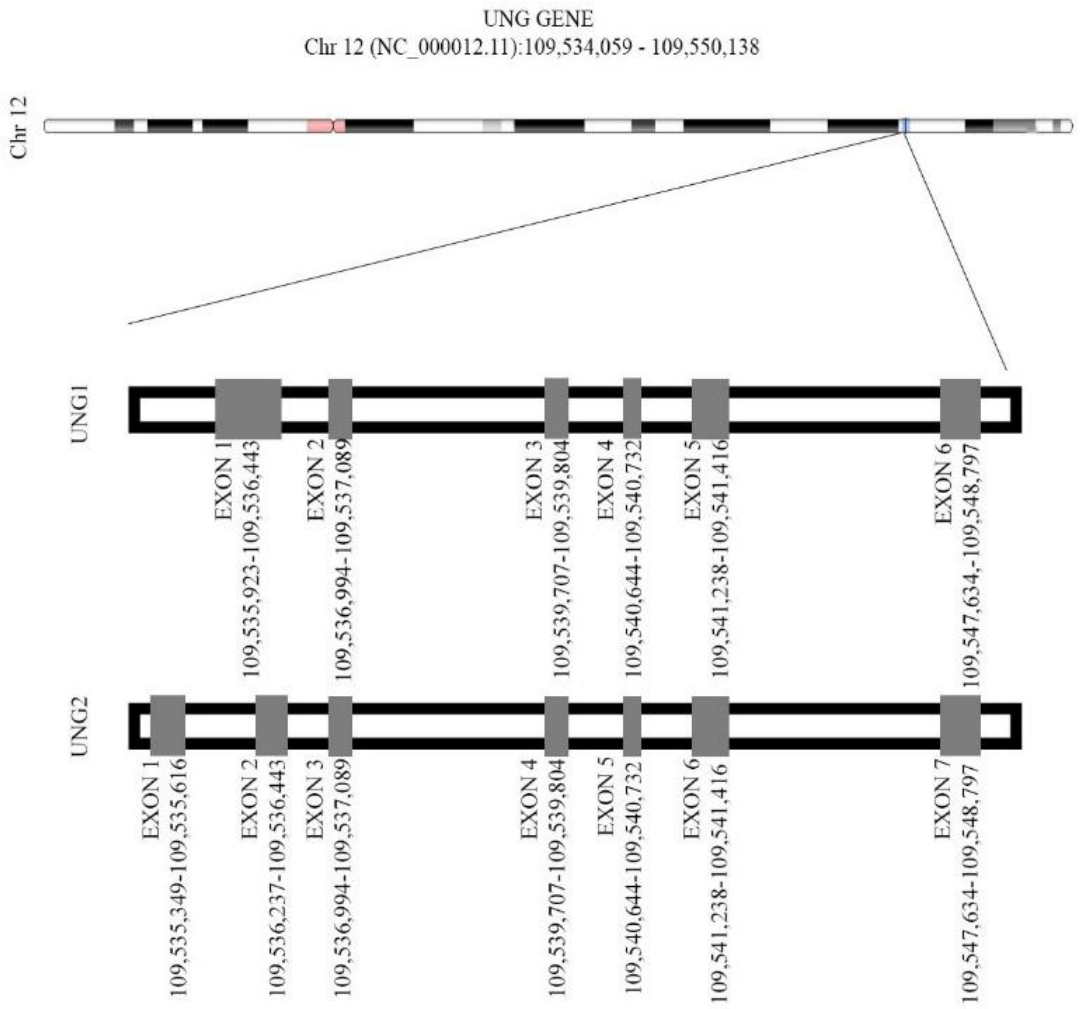


Figure 2.6: The exonic and intronic structures of *UNG1* and *UNG2* genes.

The clustering of the enzymes is provided by the sequences of the catalytic domains of the proteins. The sequences of N-terminal regions providing the localizations of the proteins are different in the different organisms. SMUG-1 is another uracil removal enzyme that presented in the *H. sapiens*. SMUG-1 is also monofunctional DNA glycosylase and a member family-3 of UNGs. The SMUG-1 enzyme acts like UDG enzyme and used for the backup mechanism of uracil removing activity in BER (75,76).

UNG1 and *UNG2* have highly similar sequences, and both of them carry 269 same amino acids in their catalytic domains, but their N-terminal regions have different sequences for the subcellular localizations (UNG1 have 35 amino acids, and UNG2 have 44 amino acids) (5,77). UDG enzymes recognize the uracil adducts by

its own catalytically active domain, and this domain is hydrolyzed by N- glycosidic bond between damaged bases and sugar moieties. This process produces abasic strands in DNA (Figure 2.7) (68). N-terminal sequences of the enzymes are important as catalytically active domains of the protein for the functionality of the UDG enzymes (78).

A B

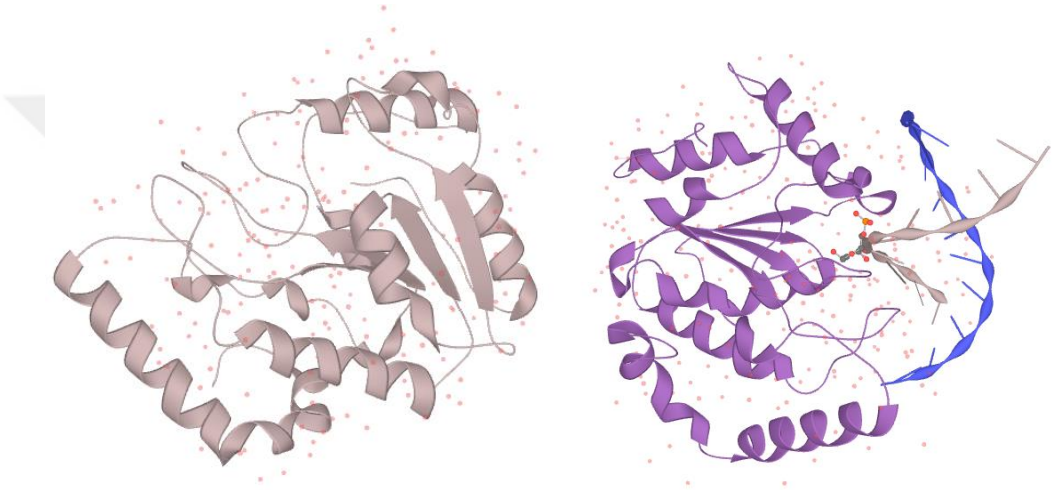


Figure 2.7: 3D Structure of UDG protein's. A. The single chain protein is folded with kinetically. B. The interaction between UDG and DNA base starts with this reaction (79,80).

In recent studies, *UNG* gene lacked organisms have been generated by molecular biological techniques in the *UNG* defected mice. Studies have been shown that back-up systems have been activated in order to maintain the uracil removal mechanism. So, insufficient or deficient production of UDG enzyme has not been caused to the failure of the organisms but, the accumulation of uracil bases was observed in the organisms' genome (81). Also, uracil removal mechanism components are attendant to the maintaining of immune system in the B lymphocytes. Investigations had been demonstrated that the triggering mechanism of immunologic disorders in the mammalians (76,82,83).

UDG enzymes have important roles for AD. For example, increased levels of folic acid and homocysteine cause neuronal DNA damage. This stress, also, leads to down-regulation of UDG protein's activity with increased uracil-incision activity (7,8). This effects have been investigated in the mouse models and cellular systems. The results have been indicated that increased effect of hippocampal neurodegeneration and toxic Aβ accumulation in the brain (6,8). This situation generates BER pathway damage in the organisms that observed in MCI patients' post mortem brain tissues. Decreased levels of UDG activities in either unaffected or effected brain regions have also decreased levels of BER activities. This result have been evaluated as non-pathologic brain regions also affected deformities of BER pathway (9–13). Besides, deficiency of this gene is observed at Down syndrome and immunologic damages. The mutations in *UNG* gene are collected and their functions are analyzed. Also, specific gene and specific disorder's interactions are analyzed, and obtained information is stored as database. On the other hand, no information had been found about the protein translated mutation that describes the relationship between sporadic AD and *UNG1* or *UNG2* variations (Table 2.2) (71,82,84–87). Chromosomal abnormalities that include the *UNG* gene have been demonstrated as clinical significant genomic changes. Also, pathological variations found on the *UNG*s are responsible for immune deficiency syndromes. On the other hand, partially duplication of *UNG* gene (nsv2773546) act likely benign. This gene duplication is affected by co-regulated genes' activities of *UNG*s (more than ten genes) (40). Also, experimental studies have been demonstrated that the methylation level changes in the *UNG* gene's promoter regions. If these methylation/de-methylation circles complete successfully, organisms' permanency continues, but another way, life of organism's ends (88).

Table 2.2: The functional consequences of UNG common variants. “SNP: Single-nucleotide polymorphism, CNV: Copy-number variations” (40,64,77,89–91).

Chromosomal locations	SNP number	Variation	Significance of the variation	Pathology	Variations' Type	Effected Genes	Diseases	References
chr12:109,535,412	rs886048904	C/T	Upstream of the gene	Uncertained	SNP	*UNG-1*, *UNG-2*	Immunodeficiency disorder	(64)
chr12:109,541,367	rs104894380	C/T	Missense	Pathogenic	SNP	*UNG-1*, *UNG-2*	Immunodeficiency disorder	(64,77)
chr12:109,190,922 – 109,989,416	nsv2773546	-	Duplication (798494 base)	Likely benign	CNV (gain)	Affected on >10 genes.	-	(40)
chr12:173,787 – 133,777,902	nsv2770609	-	Duplication (133604115 base)	Pathogenic	CNV (gain)	Affected on >1000 genes.	Dwarfism, morphological abnormalities on cardiovascular systems and ears, Polydactyl	(40)
chr12:173,787 – 133,777,902	nsv1363932 7 nsv2770609	-	Duplication (133604115 base)	Pathogenic	CNV (gain)	Affected on >1000 genes.	Morphological abnormalities on face and heart, syndactyl, growth retardations	(40)
chr12:103,044,333 – 111,639,805	nsv2778199	-	Deletion (8595472 base)	Likely pathogenic	CNV (loss)	Affected on >75 genes.	Growth retardations, morphological abnormalities	(40)
chr12:1 – 133,851,895		-	Duplication (133851895 base)	Pathogenic	CNV (gain)	Affected on >1000 genes.	Abnormalities on cardiovascular systems and face, cleft palate	(89,90)
chr12:91,438,095 - 109571015	nsv492238	-	Duplication (18132920 base)	Pathogenic	CNV (gain)	Affected on >1360 genes.	Different types of seizures, Dwarfism, hypothyroidism	(40)
chr12:282,465 – 133,773,393	nsv917029	-	Duplication (13349028 base)	Pathogenic	CNV (gain)	Affected on >150 genes.	Sacral dimple, Abnormalities on cardiovascular systems, hands feet and head	(40)
chr12:105,628,455 – 112,632,490	nsv577404 nsv529400	-	Deletion (7004035 base)	Pathogenic	CNV (loss)	Affected on >90 genes.	Speaking delays	(91)

3. MATERIALS and METHODS

The methodological research steps used in this thesis study are summarized in figure 3.1, and explained in detail in following sections.

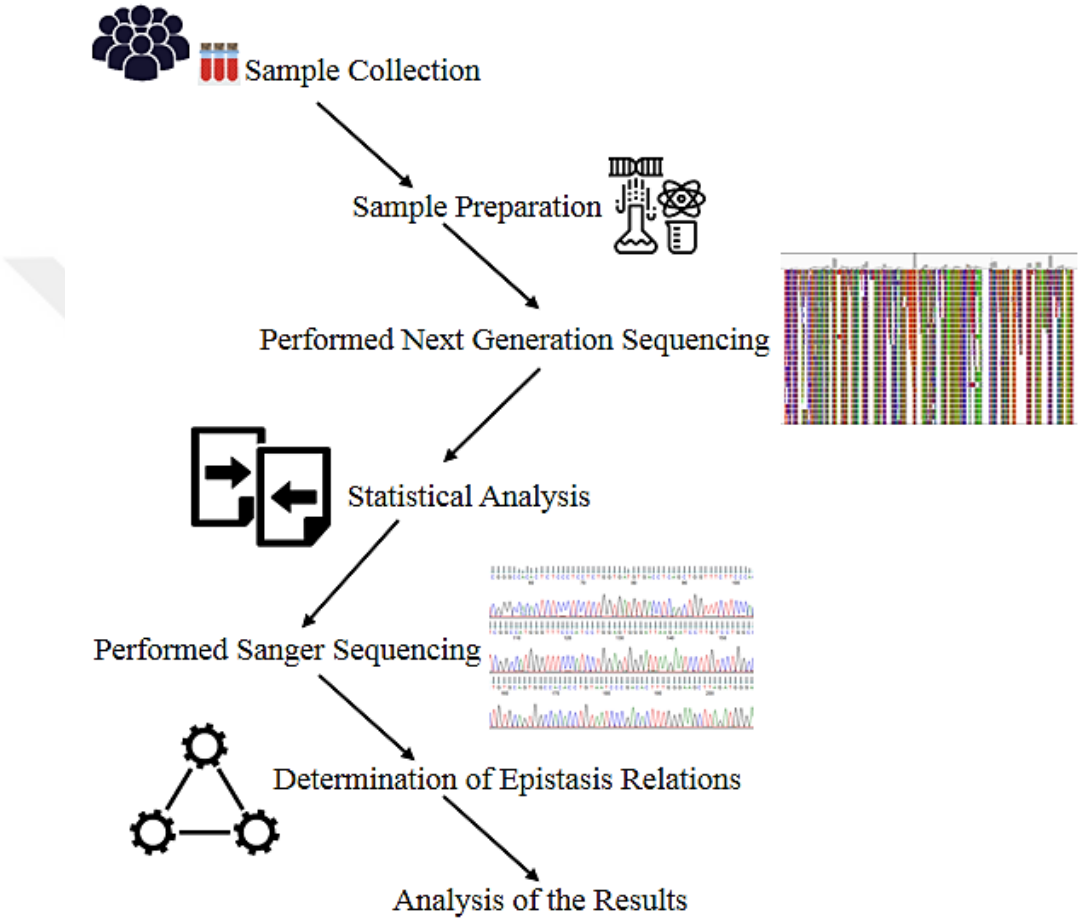


Figure 3.1: Methodological research steps.

3.1.Selection of Case and Control Participants

Patients and controls were assembled from “Clinics of Neurological Disorders at Medeniyet University’s Education and Research Hospital, Istanbul”. In this study, 221 sporadic AD patients (mean age 79.7±7.71) and 120 controls (mean age 75.2±7.21) were used. Standard psychological and physiologic examinations were performed for all the participants, and accompanying disorders were determined

(22). Cancer patients and other neurological disorder's carriers were excluded from the study. Written informed consents were obtained from all participants prior to participation in this study. The study was approved by the "Ethical Committee of Acibadem Mehmet Ali Aydinlar University and Acibadem Health Institutions Medical Research" (ATADEK 2014-700).

3.2. Isolation of DNA

Blood samples were collected in the "BD Vacutainer Blood Collection Tubes" that contained "EDTA (Becton, Dickinson and Company, USA)". DNA isolation from blood samples was performed using "DNA Mini-Blood kit (Qiagen, Germany)" according to manufacturer's procedure.

DNA samples were quantified using "NanoDrop 2000c Spectrophotometer" and "Qubit 2.0 Fluorometer (Thermo Fisher Scientific, USA)". For "Qubit 2.0 Fluorometer (Thermo Fisher Scientific, USA)" quantification the working solution was prepared by mixing "199 µL Qubit® dsDNA BR reactive" and "1 µl Qubit® dsDNA BR buffer". Ten microliters of "standard 1" and "standard 2" were mixed with "190 µl working solution", incubated for two minutes at room temperature and standard curve was generated with quantification of standards. Two microliters of DNA samples were mixed with "198 µl working solution". Mixtures were incubated for two minutes at room temperature and quantified. "The ratio of 260/280 (1.8-2.0 nm) and 260/230 (2-2.2 nm)" of DNA samples were used. DNA samples were stored at -20 C.

3.3. "Ion Personal Genome Machine (Ion-PGM) System" for targeted UNG Gene Sequencing

3.3.1. Designing of the sequencing primers

UNG gene primer ("GRCh37-hg19 human reference genome") was designed using "Ion Ampliseq Designer software (https://www.ampliseq.com)" (Figure 3.2). The designed primer covers 94.76 % of *UNG* gene region (promoter, exon and intron) and contains 54 amplicons which are divided into two primer pools. The amplicons are between 125-375 bp and the average amplicon length is 268 (±67.3)

bp (Table 3.1). The genomic regions of the designed primer are shown on the map of *UNG* gene structure (Figure 3.2).

Table 3.1: Information about targeted-NGS primers'.

Gene	Genome	Material Type	Beginning of the sequenced gene	End of the sequenced gene	Length of the Target region	Amplified region	Amplicon count	Percentage of the covered genomic region
UNG	hg19	Genomic DNA	109534879	109548796	13919 bp	13190 bp	54	94.76 %

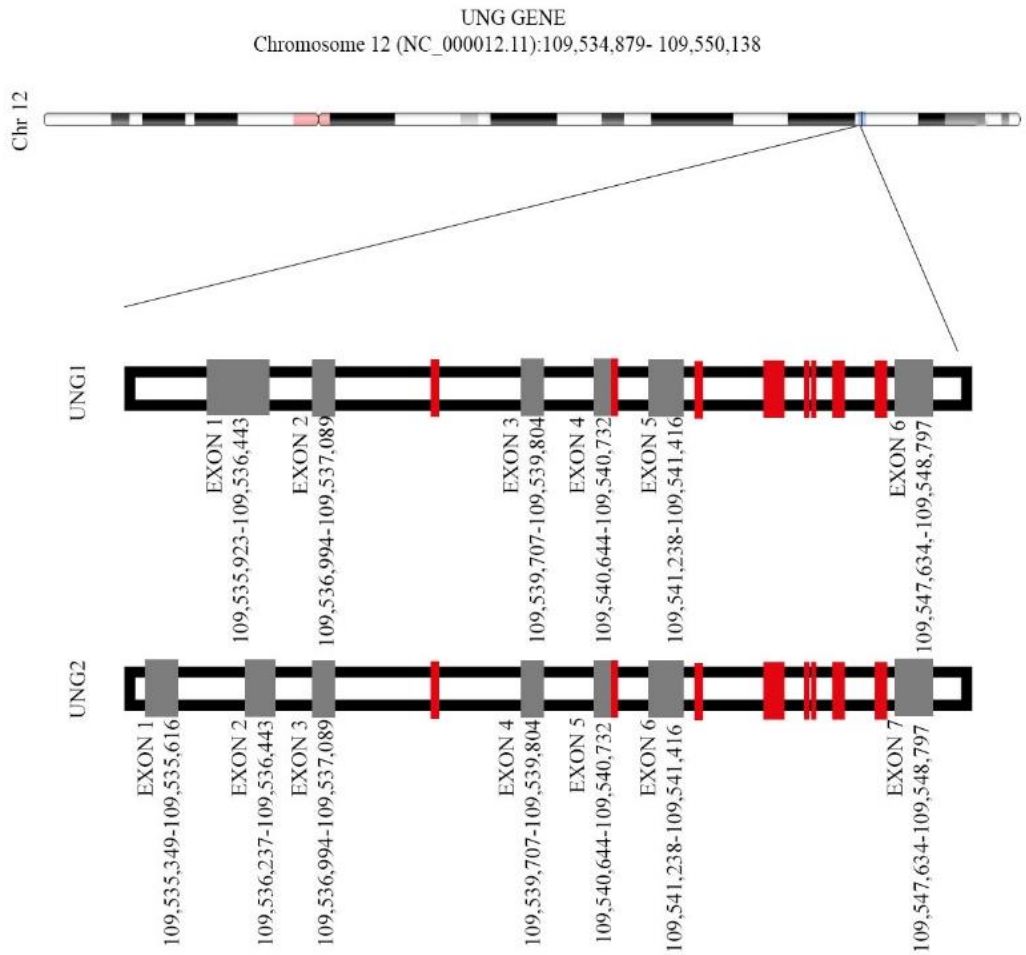


Figure 3.2: "*UNG* gene structure". Red box: uncovered regions

3.3.2. Library preparation

The steps for the library preparations are summarized in Figure 3.3.

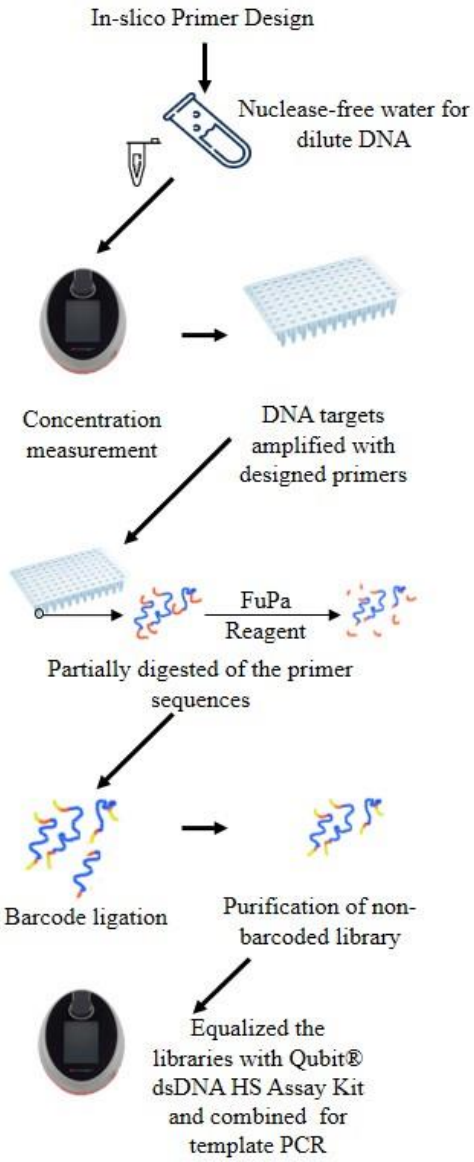


Figure 3.3: The preparation of the libraries.

3.3.2.1.Multiplex-PCR

PCR reactions for each DNA sample were prepared according to Table 3.2 in the 96-well plate and reaction was started with each primer pool (Table. 3.1). Table 3.3 demonstrates the multiplex PCR conditions that were performed using "Verity Thermal Cycler (Applied Biosystems, USA)".

Table 3.2: Final concentrations of the "multiplex PCR reactions".

Components	Final Conc.
"Ion-Ampli Seq Hi-Fi Mix"	1 x
"Ion-Ampli Seq Primers"	1 x
DNA	20 ng

Table 3.3: The multiplex PCR thermal-cycle conditions.

Stages	Steps	Temperature's	Duration
Incubation	"Enzyme activation"	99°C	2 min.
22 x	"Denaturation"	99°C	15 sec.
	"Annealing and extension"	60°C	4 min.
	-	10°C	Hold

3.3.2.2.Partial digestion of the PCR products

The PCR products were digested and phosphorylated with "2 µl FuPa enzyme" according to the thermal cycler conditions in Table 3.4.

Table 3.4: Incubation conditions of "Fu-Pa Reagent".

Temperatures	Duration
50°C	10 min.
55°C	10 min.
60°C	20 min.
10°C Incubation	Up to 1 hours

3.3.2.3.Ligation of barcodes and adaptors

The “Ion Xpress barcode”s were diluted in 1:4 ratios using 2 µL “Ion Xpress barcodes”, 2 µl “Ion P1 Adapters” and 4 µL nuclease free H_2O. Two microliters of diluted barcodes, “4 µl Switch Solution” and “2 µl DNA ligase” were added into digested PCR products. Then, the barcode ligations were performed according to the thermal cycler conditions in Table 3.5.

Table 3.5: Barcode ligation conditions.

Temperatures	Times
22°C	30 min.
72°C	10 min.
10°C	Up to 1 hours

3.3.2.5. Purification of the barcoded library

“Agencourt AMPure XP Reagent (Beckman Coulter, USA)” was brought to room temperature and mixed thoroughly with vortex. After ligation step 45 µl “AMPure XP Reagent (Beckman Coulter, USA)” was added to each library and mixed thoroughly. The mixture was incubated for 5 minutes at room temperature and for 2 minutes on “DynaMag-96 Magnet”. The supernatants were discarded, and the beads were washed with “freshly prepared 150 µl ethanol” for 2 times. Ethanol was removed, and beads were air dried for 5 minutes at room temperature.

3.3.2.6. Library equalizing

Purified libraries were amplified with “50 µl Platinum PCR SuperMix High-Fidelity” and “2 µl Library Amplification Primer Mix” according to thermal cycler conditions at Table 3.6. Amplified libraries were purified with a two-round purification step. First “25 µl Agencourt AMPure XP Reagent (Beckman Coulter, USA)” was added on amplified libraries and the mixture was incubated for 5 minutes at room temperature and for 5 minutes on “DynaMag-96 Magnet”. Second purification round was performed with “60 µl Agencourt AMPure XP Reagent

(Beckman Coulter, USA)". The incubation steps were same with first round of purification. Beads were washed as told in method 3.3.2.5.

Table 3.6: "The thermal-cycle conditions of library amplification's".

Stages	Temperatures	Times
Incubation	98°C	2 min.
5x	98°C	15 min.
	64°C	1 min.
Incubation	10°C	Hold

Fifty microliter low TE was used for resuspension of bead pellets. Resuspended pellets were placed on "DynaMag-96 Magnet" and concentration of libraries was quantified with "Qubit 2.0 Fluorometer". Working solutions and standards were prepared as told in method 3.2. Five microliters from libraries were mixed with "195 µl working solutions". Dilution factors were calculated according to quantification results and libraries were equalized to 100 pM. Ten microliters from equalized libraries were mixed.

3.3.3. "Ion Sphere™ Particles (ISPs)" positive template preparation

Emulsion PCR based template preparation steps were summarized in Figure 3.4.

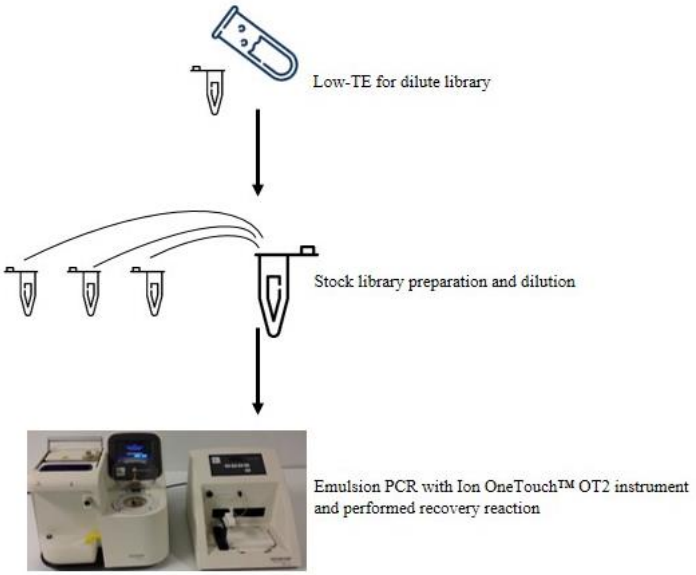


Figure 3.4: The preparation of enrichment.

3.3.3.1. Initializing the “Ion One Touch 2 (OT2) Instrument”

The instrument was turned on and the lid opened. One hundred and fifty microliters “Ion OneTouch™ Breaking Solution” containing recovery tubes were placed into their centrifuge slot. “Recovery router” was placed in center slot of centrifuge (Figure 3.5). Amplification plate was inserted into the heat block as summarized in figure 3.6. One of the reagent tubes was half-filled with oil and the other reagent tube was quarter-filled with recovery solution. The recovery tubes were inserted into their ports (Figure 3.7).

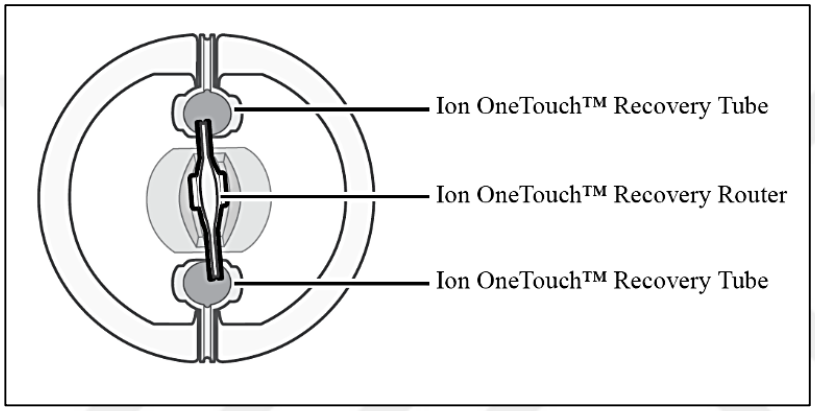


Figure 3.5: “Ion One-Touch Instrument”s centrifuge router schema (“figure was illustrated from Ion-PGM Hi‑Q OT2 Kit’s protocol”).

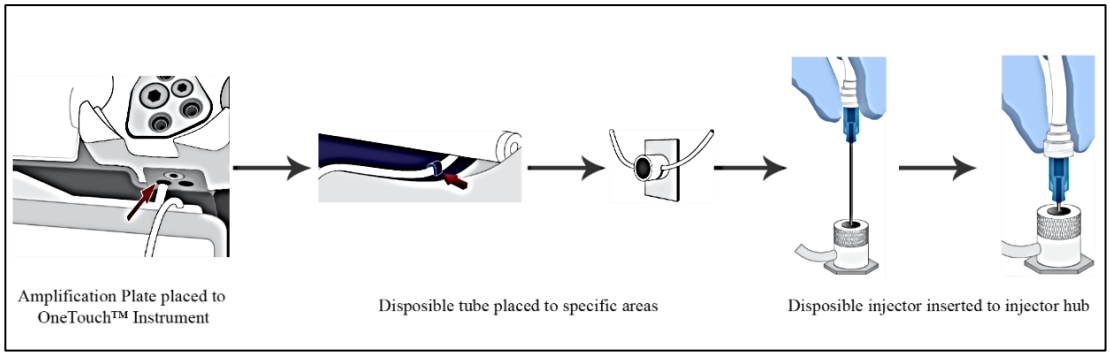


Figure 3.6: Placed the “amplification plate” to “Ion One-Touch Instrument” (“figure was illustrated from Ion-PGM Hi‑Q OT2 Kit’s protocol”).

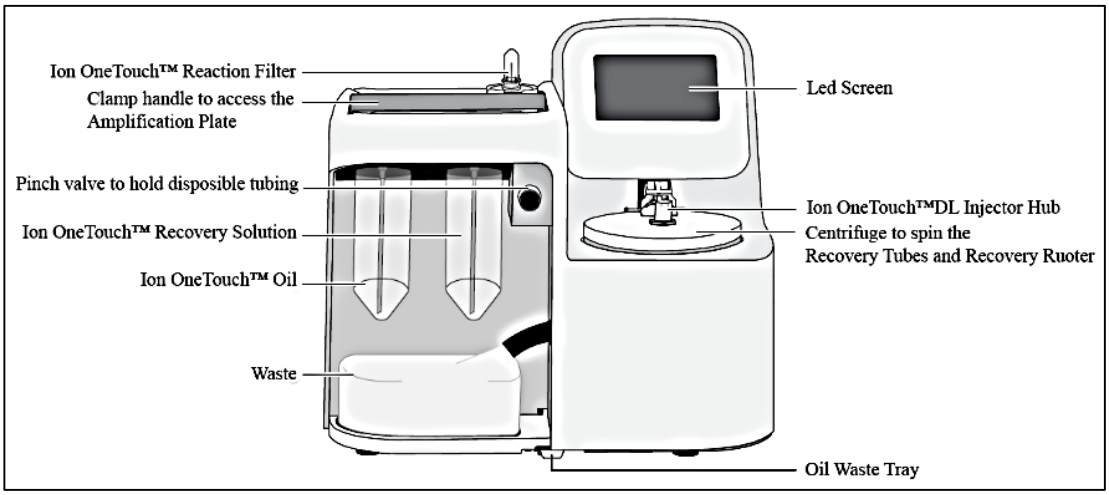


Figure 3.7: "Ion One-Touch Instrument"s schema ("figure was illustrated from Ion-PGM Hi‑Q OT2 Kit's protocol").

3.3.3.2.Preparation and installation of amplification solution

"Ion-PGM™ HiQ™ Reagent Mix" and "Ion-PGM™ HiQ™ ISPs" were allowed to come room temperature before using. Two microliter of mixed library from method 3.3.2.11 was added into 23 µl of nuclease free water. Amplification solution components (Table 3.7) were added into 800 µl containing "Ion-PGM™ HiQ™ Reagent Mix tube".

Table 3.7: Amplification solution components.

Reagents	**Volumes**
Nuclease free H_2O	25 µl
"Ion-PGM™ HiQ™ Enzyme"	50 µl
Diluted Library	25 µl
"Ion-PGM™ HiQ™ ISPs"	100 µl

Complete amplification solution was vortexed for 5 seconds and centrifuged for 2 seconds. After centrifugation complete amplification solution was loaded into "Ion One-Touch Reaction Filter" from sample port (Figure 3.8). "1.7 ml Ion One-Touch Oil" was slowly added into "One-Touch Reaction Filter". Loaded "Ion One-Touch Reaction Filter" was inserted to "OT2 instrument" and "PGM: Ion PGM HiQ OT2 Kit 400" program was started.

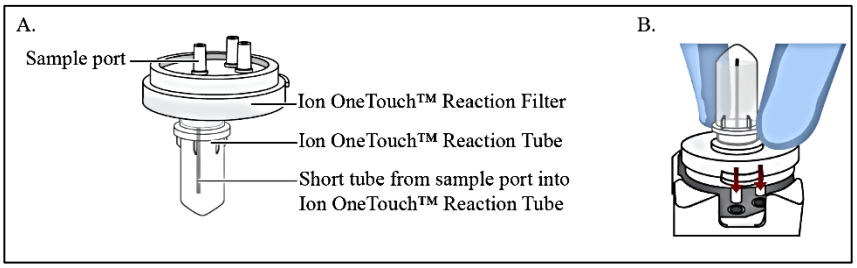


Figure 3.8: "Ion One-Touch Reaction Filter"s schematic description. A. Filter's schema. B. Insertion of the filter's to the instrument's port ("figure was illustrated from Ion-PGM Hi‑Q OT2 Kit's protocol").

3.3.3.3. Recovering of "Template-Positive Ion-PGM Hi-Q ISPs"

At the end of the amplification, final centrifugation was performed. After centrifugation, all but 100 µl of recovery solution was removed from recovery tubes without disturbing the "ISPs pellets". "ISPs pellets" were resuspended in the remaining recovery solution and 500 µl "Ion-OneTouch Wash Solution" was added to each recovery tube. ISPs were dispersed by pipetting and transferred in a new 1.5 ml tube. ISPs were centrifuged for 2.5' at 15,500xg. All but 100 µl wash solution was removed without disturbing the pellet.

3.3.3.4 "Ion Sphere™ Particles (ISPs)" positive template enrichment

"Melt-off solution" was prepared with "280 µl Tween Solution and 40 µl 1mM NaOH". "Dynabeads MyOne Streptavidin C1 Beads" was vortexed for 30 seconds then centrifuged for 2 seconds. "13 µl Dynabeads MyOne Streptavidin C1 Beads" was transferred to a new 1.5 ml tube and incubated on "DynaMag™ magnet" for 2 minutes at room temperature. Supernatant was discarded, and beads were resuspended with "130 µl MyOne™ Beads Wash Solution". Recovered "Template-Positive Ion-PGM Hi-Q ISPs" from method 3.3.3.3, 130 µl resuspended "Dynabeads MyOne Streptavidin C1 Beads", 300 µl "Ion OneTouch Wash Solution" and 300 µl melt-off solution was loaded into "8-well strip" as shown in figure 3.9.

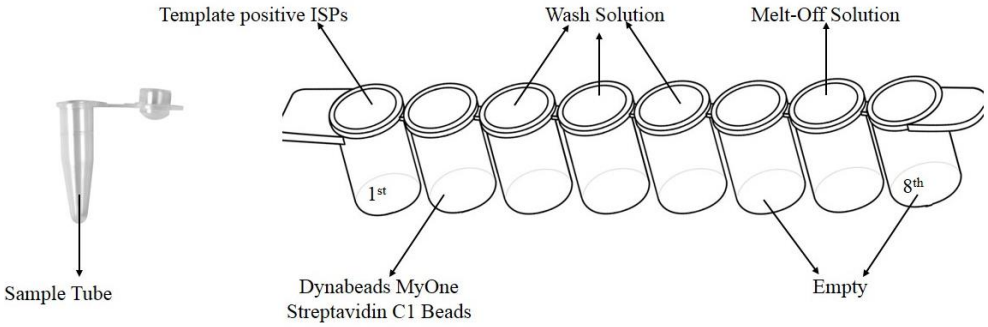


Figure 3.9: 8-well strip filling for the enrichment of "template positive Ion PGM HiQ ISPs".

New tip was placed to tip arm. Loaded 8-well strip was placed on "Ion OneTouch™ ES" and 10 µl "Neutralization Solution" containing 0.2 ml tube was placed into tip loader hole. Start/stop button was pressed for initializing the enrichment.

3.3.4. Cleaning and initializing the "Ion-PGM System"

3.3.4.1. Creation of a planned run

Torrent Suite server was used for creating a new run plan.

3.3.4.2. Preparation of gas cylinders leak test

Main tank's shutoff valve was opened. The pressure was adjusted to 30 psi and valve was closed. The monitor was observed for 5 minutes for avoiding any pressure change.

3.3.4.3. Chlorite and 18-MΩ H2O cleaning

Cleaning chip was placed to chip socket. Cleaning bottles were rinsed with 18-MΩ H_2O for two times. "Ion Cleaning Tablet" was dissolved in the 1 L of 18-MΩ H_2O and 1 ml 1 M NaOH was added into solution. The solution was filtered with 0.22 µm filter. Chloride cleaning option was selected from cleaning menu on device. Chloride cleaning bottle was filled with 250 ml solution, filled bottle was placed to W1 position and cleaning bottles were placed to W2 and W3 positions (Figure 3.10). After chloride cleaning is completed, solution bottle was removed from W1 position and the sipper was washed. 250 mL 18 MΩ containing water bottle was placed to

W1 position and water cleaning was started from device. When the cleaning was completed all bottles and washing sippers were removed from the device.

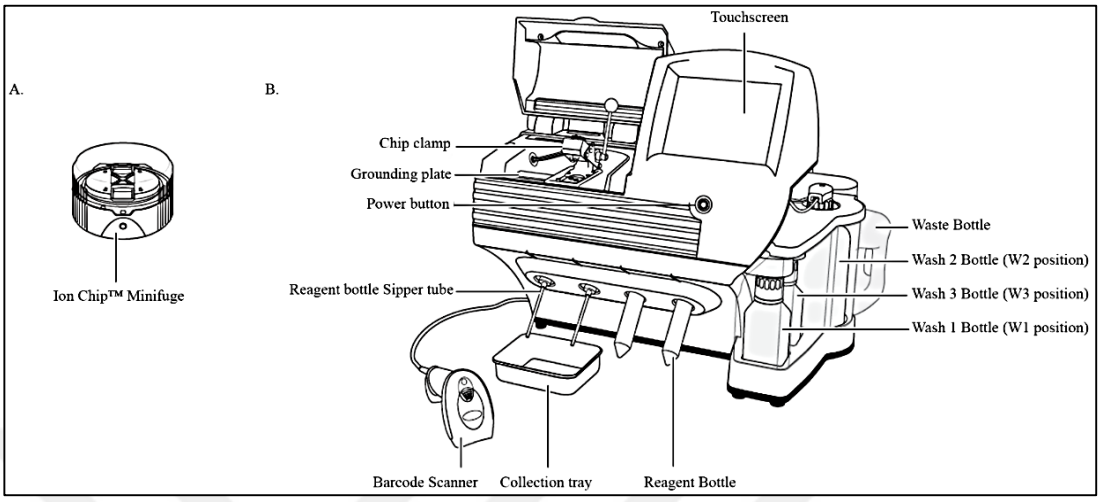


Figure 3.10: Schematic description of "Ion-PGM devices'". A. "Ion-Chip" mini centrifuge. B. "Ion-PGM device" ("figure was illustrated from Ion-PGM Hi‑Q OT2 Kit's protocol").

3.3.4.4. Initialization of the system

New sippers were placed to washing positions. "Wash 2 bottle" was rinsed with 200 ml 18-MΩ H_2O for 3 times. Two litre of 18-MΩ H_2O, entire bottle of "Ion-PGM HiQ Sequencing-W2 solution" and 70 µl 100 mM NaOH was added to the bottle. The bottle was inverted 5 times for mixing. Three hundred and fifty microliters 100 mM NaOH was added to "Wash 1 bottle". Wash 3 bottle was filled with 50 ml Hi‑Q Sequencing-W3 Solution. Initialize option was selected from device, the barcode of "Ion-PGM HiQ Sequencing-W2 solution" was scanned and initializing was started.

Device automatically adjusts it's pH between 7.49 and 7.6. When the adjustment finalized old sippers at dNTP positions were discarded, and new sippers were placed. New reagent tubes were labelled as "dGTP, dCTP, dATP, and dTTP". Twenty microliter "dNTP stock solutions" was added to labelled tubes. Tubes were placed to their own positions and tightened. Touchscreen orders were followed for completing the initialization.

3.3.4.5. Chip loading

“Semiconductor Ion 318™ Chip” was used for sequencing. “Control Ion Sphere™ Particles (ISPs)” was vortexed and centrifuged briefly. “5 µl Control ISPs” was added into enriched “ISPs positive template”. The solution was mixed thoroughly and centrifuged for 2 minutes at 15,500xg. All but 15 µl of supernatant was discarded. Pellet was resuspended with “12 µl Sequencing Primer” and incubated in a thermal cycler at 95°C for 2 minutes and at 37°C for 2 minutes. After incubation “3 µl Ion PGM™ Hi‑Q™ Sequencing Polymerase” was added into tube and incubated at room temperature for 5 minutes.

During polymerase incubation chip check was initialized with cleaning chip, and new chip was placed to chip socket when prompted. The barcode of new chip was scanned. After chip check was finalized all liquid inside the chip was discarded.

Chip was placed in the bucket of “Ion Chip™ Minifuge”. ISPs were loaded into chip and chip was centrifuged for 30 seconds chip tab pointing inside and for seconds chip tab pointing outside. The ISPs was pipetted up and down one time and removed from the chip. Chip was turned upside-down and centrifuged for 5 seconds. All the remaining liquid was discarded from the chip.

Chip was placed into chip socket and sequencing was started with pre-planned run (92).

3.4.Analyzing of the NGS Results

Bioinformatics analysis of the results was performed with “Ion-Reporter Software” and “CLC Genomics Workbench”.

3.4.1. Analyzing of the results: “Ion Reporter Software”

Raw data from targeted gene sequencing was processed by “Torrent Suite Software”s automated plugins. For bam files, FASTQ files were aligned to “GRCh37-hg19 human genome reference”. BAM files were processed by “variant caller plugin”, and VCF files were created. BAM and VCF files were transferred to “Ion Reporter Software”. VCF files were analyzed according to their location, zygosity, position, type, and accession number of the variations.

3.4.2. Analyzing of the results: "CLC Genomics Workbench"

VCF files from "Ion Reporter Software" were used in "CLC Genomics Workbench (9.0.1., Qiagen, USA)". "Fisher-Exact Test" was performed for comparative analyses of case-control groups. For analyzing all variants in case-control groups, p value was set to 0.9 as maximum threshold and "Bonferroni Correction" was applied. Results were sorted according to their p-values for determining significantly important variations. In addition, variants showing low frequency in sporadic AD group were excluded.

3.5. Selection of the significantly important variations

Significantly important variations were selected according to their p values, location and their relation with other variations. Significantly important variations were selected for confirmation with Sanger sequencing.

3.6. Significantly Important Variations Confirmation using Sanger Sequencing

3.6.1. Primer design

Primers were designed using "Primer-Blast (NCBI) designing tool". Designed primers are shown in Table 3.8.

Table 3.8: Sanger primers that covered selected variations.

Primer's ID	Covered Genomic Locations		Sequences of primers
U2-V2	chr12:109,535,699 - 109,536,033	Forward	5'-GCG CCT CTG ACT CGG TAA AC-3'
		Reverse	5'-CGT GGT CTT TTC GCG GCA-3'
U3	chr12:109,540,012 - 109,540,432	Forward	5'-ATC CCT GGC CTC TAC CAC-3'
		Reverse	5'-GGA TTT AAC AGC ACT CTG GA-3'
U4	chr12:109,537,347 - 109,537,699	Forward	5'-TCC CGC GGT GTC AGA TGT T-3'
		Reverse	5'-ACC AGG TTT CAT TGG TCT TGA ACT C-3'
U5	chr12:109,545,726 - 109,546,164	Forward	5'-GGC ACA AAA TGG AAA GTA AAT GGC-3'
		Reverse	5'-TCC CAG CCA AGA GAC CCT G-3'
U5.1	chr12:109,546,035 - 109,546,370	Forward	5'-GTA CAA CTG GTC CCT GGC TTA-3'
		Reverse	5'-CCA CCT CCT AGG CTC AAG TTA C-3'
U6	chr12:109,546,660 – 109,547,130	Forward	5'-CTC TTG GTC CCA TGC AGT T-3'
		Reverse	5'-TGC AAA GAG CAC CAC CCT C-3'
U7	chr12:109,548,405- 109,548,703	Forward	5'-TTG TTA TTC CTT GGG TGT TGG-3'
		Reverse	5'-GGT GCA GAC CCA ATA AAG-3'
U8	chr12:109,540,719 - 109,541,202	Forward	5'-GTG GCC CAA GCA AGG TAA-3'
		Reverse	5'-AGG CAA ATC AGC CCA ATA CC-3'
U10	chr12:109,541,181 - 109,541,640	Forward	5'-CTG GGT ATT GGG CTG ATT TGC-3'
		Reverse	5'-CCA GTT GCT TTG TCC TC-3'

3.6.2. "Polymerase Chain Reaction (PCR)" and purification of PCR products

PCR products were amplified using 50-100 ng DNA, "GeneAmp dNTP Blend (Thermo Fisher, USA)" and "Taq DNA Polymerase (Invitrogen, USA)". PCR components and conditions were shown in Table 3.9.

Table 3.9: The amplification conditions for the target sequences. A. PCR ingredients for "Sanger sequencing" samples. B. Thermal-cycle conditions.

A

PCR Ingredients	The Final Concentrations
Nuclease free H_2O	to 25 µl
"PCR-Buffer"	1 x
"$MgCl_2$ (50 mM)"	1 mM (U1, U2, U5, U6, U8, U9, U10 primers) 1.5 mM (U3, U4 primers)
"Primers (forward and reverse)"	0.5 mM
"Taq-DNA Polymerase"	2 Unit
Total vol. of the reaction	25 µl

B

Step		Temperatures'	Times for the reaction
"Initial-denaturation"		94°C	5 min.
35x	"Denaturation"	94°C	45 sec.
	"Annealing"	58.5°C (for U1, U2, U4, U6) 60°C (for U3, U4, U5, U8, U9, U10)	30 sec.
	"Extension"	72°C	30 sec. (U1, U2 primers) 60 sec. (U3, U4, U5, U6, U8, U9, U10 primers)
"Final-extension"		72°C	10 min
Hold at		4°C	∞

PCR products were separated on "1.5% agarose gel for 1.5 h at 100 V". EtBr was added in gel and obtained bands were visualized by "ChemiDoc XRS+ Imaging System (Bio-Rad)".

Two microliters of "ExoSAP-IT™ (Applied Biosystems, USA)" was mixed with 5 µl PCR products for purification. The mixture was incubated at "37° C for 15 min and at 80° C for 15 min".

3.6.3. Cycle sequencing

"Big-Dye Terminator v3.1 Cycle-Sequencing (Thermo Fisher Scientific, USA)" kit was used for de novo sequencing. Sequencing reaction was prepared using "2 µl Terminator 3.1 Ready Reaction Mix, 2 µl BigDye™ Terminator v3.1 5X Sequencing Buffer, 2 µl forward or reverse primer 1.5 µl PCR product and 2.5 µl deionized

water". Sequencing was performed according to the thermal cycler conditions in Table 3.10.

Table 3.10: The thermal-cycle conditions for "Big-Dye reaction".

Step		Temperatures'	Times for the reactions
"Initial Denaturation"		96°C	1 min.
25x	"Denaturation"	96°C	10 sec.
	"Annealing"	50°C	5 sec.
	"Extension"	60°C	4 min.
Hold at		+4°C	To one week

3.6.4. Cleaning of the residual labels

"Big-Dye XTerminator Purification Kit (Thermo Fisher Scientific, USA)" was used for purification of cycle sequencing products. The products were mixed with "45 µl SAM Soluıtion and 10 µl XTerminator solution". Mixture was "vortexed for 30 min at 2000 rpm and centrifuged for 5 min at 1000xg". Supernatant was transferred to a new 0.2 ml PCR tube.

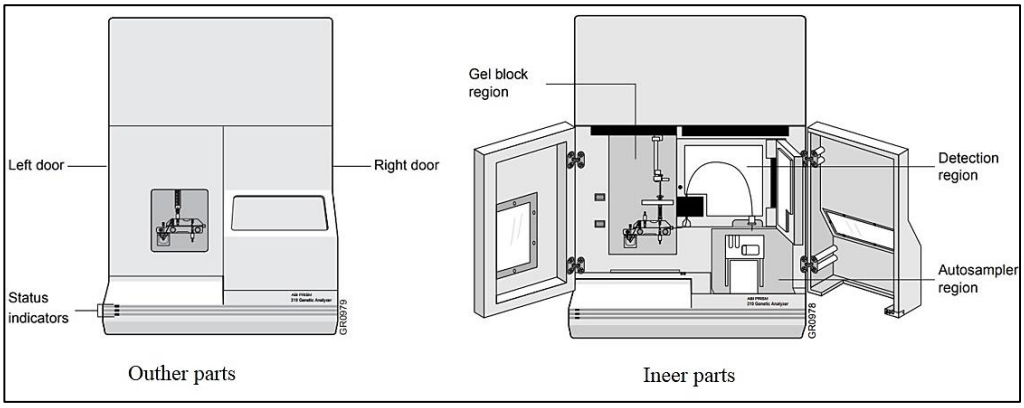
3.6.5. Sanger sequencing

Sanger sequencing was performed with "ABI-PRISM 310 Genetic Analyzer and AppliedBiosystems 3500Dx Genetic Analyzer (Thermo Fisher, USA)" (Figure 3.10).

3.7. Analysis of the Sanger results

Abi files from sequences were analyzed using "Finch TV (1.4., Geospinoza) and Chromas Pro (2.4.1., Technelysium)" analyzing tools.

A



B

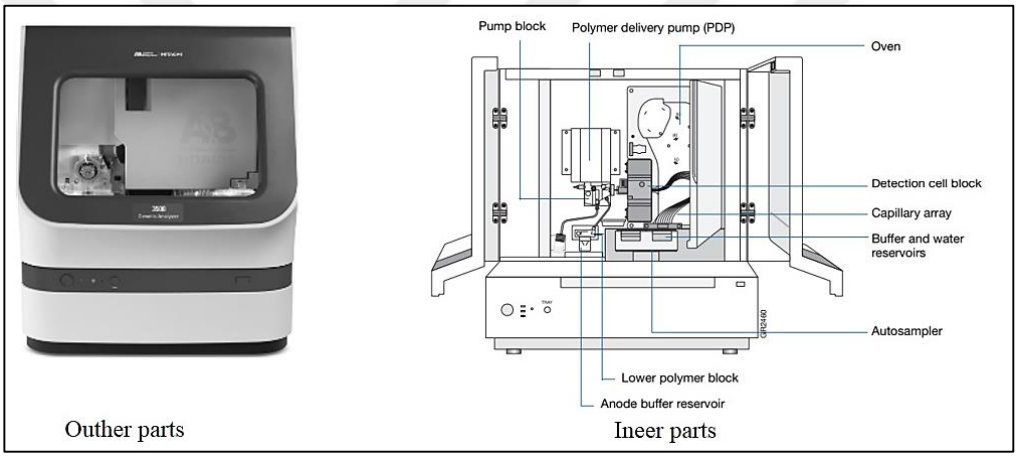


Figure 3.11: The instruments of Sanger sequencing. Interior and exterior parts are described in the illustrations. A. “ABIPRISM 310” is one capillary system for performing of Sanger sequencing. B. “Applied Biosystems-3500Dx” is eight capillary system for performing of Sanger sequencing.

4. RESULTS

4.1.Sample Selection

In this thesis study, 221 sporadic AD patients (mean age 79.9±7.71) and 120 controls (mean age 75.2±7.21) were used. Average ages of sporadic AD patients and control groups are similar (Figure 4.1). Standard psychological and physiologic examinations were performed for all the participants, and accompanying disorders were determined (22). Cancer patients and other neurological disorder's carriers were excluded from the study. MMSI scoring of sporadic AD patients indicating the severity of the disease are shown in Table 4.1.

Table 4.1: The clinic characteristics of the participants.

Characteristics of the group	**Participants**	
	Cases	**Controls**
The total count of the participants	221	120
Gender		
Females (percentage and count)	58.6% (129)	52,9 % (62)
Males (percentage and count)	41.4% (93)	47,1 % (58)
Average of the age (standard deviation)	79.7 (±7.71)	75.2 (±7.21)
Grades of the disorders' (with MMSI scoring)		
Severe sporadic AD (percentage and count)	34.7% (77)	-
Moderate sporadic AD (percentage and count)	43.2% (96)	-
Mild sporadic AD (percentage and count)	22.1% (49)	-

4.2.Analysis of targeted *UNG* gene sequencing results

In this thesis study, *UNG* gene was sequenced in 198 sporadic AD patient and 50 control samples using "Ion-PGM system". "Ion-Reporter tool (Thermo Fisher Scientific, USA)" and "CLC Genomics Workbench 9.0.1 (Qiagen, USA)" were used for analysis.

4.2.1. Quality of NGS data

The results of "Ion-PGM system" were trimmed with the qualified standards of the system and aligned to "GRCh37 (2009, February)". The good quality of the alignment (99% aligned bases) is shown in Figure 4.1. The VCF files with GRCh37 annotations in "Ion Reporter™ Software" were used in "CLC Genomics Workbench". In addition, the integrity of the sequenced amplicons was analyzed with IGV tool.

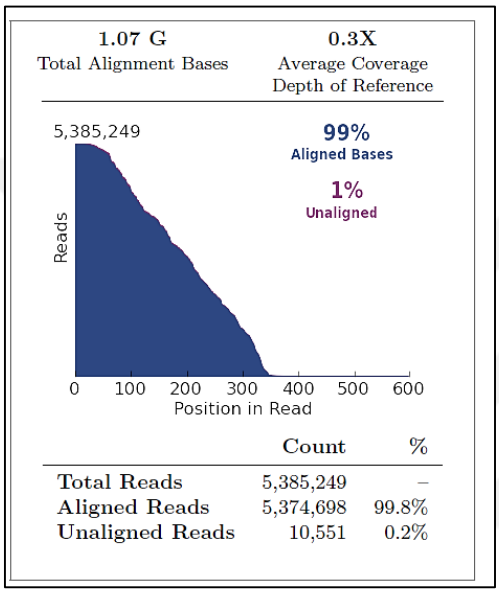


Figure 4.1: Aligned all bases and the depth of the average coverage.

To assess the quality of the libraries sequenced, the basic quality statistics for Ion Torrent datasets were determined using "CLC genomics workbench software". For example, quality distribution shows that more than 95% reads have average "PHRED quality scores (Q score)" is ≥20 (Figure 4.2) with no ambiguous bases and GC content is around 50 (Figure 4.3). Q scores measure "the base calling accuracy" that is the probability of incorrect base call. If Q score is 20, the probability of an incorrect base calls 1 in 100 times and if Q score is 30 it is 1 in 1000 times. Q scores are ≥20 accepted as a good quality reading in NGS.

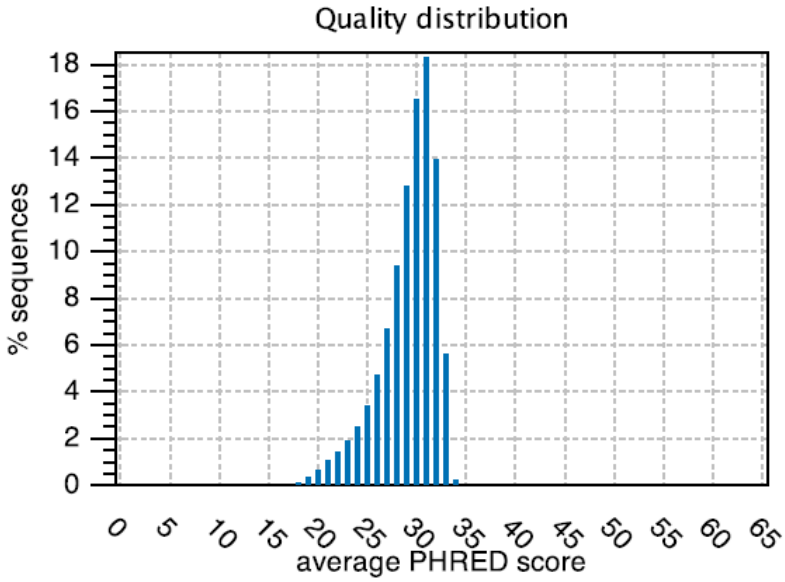
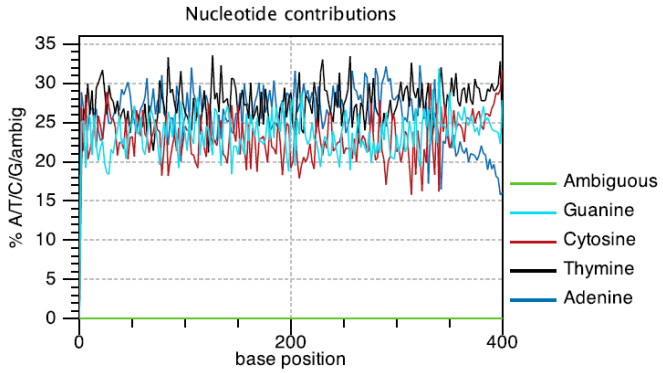


Figure 4.2: “PHRED quality score distribution”. The distribution of average “PHRED quality score” is plotted on X-axes and percentage of sequences on Y-axes for “Ion Torrent data”.

A



B

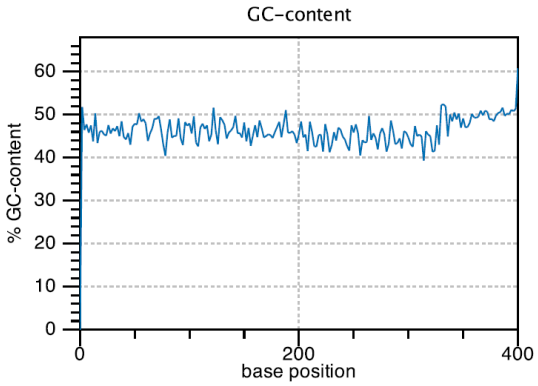


Figure 4.3: A. Nucleotide contributions. B. GC-content

4.2.2. Data analysis with "Fisher Exact Test" in "CLC Genomics Workbench"

A comparative analysis of sporadic AD patient-control groups was performed by applying the Fisher exact test at the "CLC Genomics Workbench" using the vcf files generated by "Ion Suit software". The "Fisher Exact Test" was used to identify variants that are significantly more common in the sporadic AD patient samples than in the control samples. In order to be able to examine all variants in the patient and control groups, a maximum threshold of p value (p=0.9) was used and only variants below p value of 0.9 was reported. In addition, "Bonferroni correction" was applied to the analysis. To determine statistically significant variants between the variants of the patient and control groups, the Fisher exact test results were ordered according to p value (Table 4.2). All 26 variants in Table 4.2 were selected for the "Sanger sequencing confirmation".

Table 4.2. Comparative analysis of genetic variants of sporadic AD patient and control groups by “Fisher's exact test” and ranked by p value.

UNG gene, chromosome 12	Type	dbSNP (rs)	Ref	Allele	Genotype	Intron/ exon	Zygosity	P value Fisher exact test	AD Count	Control Count	MAF
109,540,276	SNP	rs79744525	G	A	G/A	Intronic	Heterozygous	0.0147	18/198	0/50	-
109,537,423^ 109,537,424	INDEL	rs1610925	T	TTA	TTA/TTA	Intronic	Homozygous	0.0201	6/198	0/50	0.2450
					T/TTA		Heterozygous		61/198	9/50	
109,541,043	SNP	-	A	G	A/G	Intronic	Heterozygous	0.0239	16/198	0/50	-
109,548,571^ 109,548,572	INDEL	-	G	GT	G/GT	3'-UTR (exon 7)	Heterozygous	0.0261	23/198	1/50	-
109,537,562	SNP	rs80001089	T	G	T/G	Intronic	Heterozygous	0.0486	41/198	5/50	-
					G/G		Homozygous		1/198	0/50	
109,545,819	SNP	-	A	T	A/T	Intronic	Heterozygous	0.0626	7/198	0/50	-
	Deletion			-	TAAAAAA/ TAAAAA			0.7984	4/198	0/50	
109,546,886	SNP	rs2268406	T	G	G/G	Intronic	Homozygous	0.0491	4/198	0/50	0.142971
					T/G		Heterozygous		57/198	9/50	
109,535,809	SNP	rs1018782	A	G	A/G	Intronic	Heterozygous	0.0948	54/198	9/50	0.20627
					G/G		Homozygous		2/198	0/50	
109,541,051.. 109,541,052	MNV	-	AC	GT	AC/GT	Intronic	Heterozygous	0.1269	9/198	0/50	-
109,541,435 ^109,541,436	INDEL	rs144108564	C	CT	C/CT	Intronic	Heterozygous	0.1269	8/198	0/50	0.0648
					CTTTTTTTTC /CTTTTTTTT TCT				1/198	0/50	
109,545,820	INDEL	-	AT	TA	T/TTA	Intronic	Heterozygous	0.1288	15/198	1/50	-
109,541,007	SNP	-	C	T	C/T	Intronic	Heterozygous	0.1603	8/198	0/50	-
109,541,462^ 109,541,463	INDEL	-	T	A	T/TA	Intronic	Heterozygous	0.1603	8/198	0/50	-
109,541,009	SNP	-	A	C	GAC/GCC	Intronic	Heterozygous	0.1719	2/198	2/50	-
					A/C				24/198	3/50	

Continuous

Location	Type	rs number	Ref	Allele	Genotype	Intron /exon	Zygosity	P-value Fisher exact test	AD Count	Control Count	MAF Homo sapiens
109,541,004	SNP	-	G	A	G/A		Heterozygous	0.2023	7/198	0/50	-
109,545,951	SNP	rs2430678	A	G	A/G	Intronic	Heterozygous	0.2023	7/198	0/50	0.045927
109,546,114	SNP	rs2515913	C	T	C/T		Heterozygous	0.2023	7/198	0/50	0.040136
109,535,857	SNP	rs1018783	T	A	A/A	Intronic	Homozygous	0.2081	4/198	0/50	0.194688
							Heterozygous		45/198	9/50	
109,541,008	MNV	-	A	C	GA/AC	Intronic	Heterozygous	0.6514	7/198	0/50	-

"SNP: Single nucleotide polymorphism, MNV: Minor nucleotide variant, MAF: Minor allele frequency. INDEL: Insertion or deletion"

4.3. Confirming NGS results with Sanger sequencing

All NGS technologies provide base-call, insertions and deletions etc. errors in the sequenced reads and thus NGS data is error-prone. Therefore, Sanger sequencing confirmation is necessary for NGS variants. In this study, Sanger sequencing confirmation showed that 12 *UNG* gene variants identified from "Ion-PGM sequencing" experiments are false positive (Table 4.3 and Figures 4.4 - 4.10.), and 6 UNG variants are positive (Table 4.4 and Figures 4.12 - 4.22).

Table 4.3: The false positive results.

Primer's ID	Location, chromosome 12	rs number	Genotype	Intron/ exon	Fisher exact test p-value	MAF
U3	109,540,276	rs79744525	G/A	Intronic	0.0147	-
U5	109,545,819	-	A/T AT/TA	intronic	0.0626 0.7984	-
	109,545,820	-	T/TTA AT/ATTA	intronic	0.1288	-
U7	109,548,571 ^109,548,572	-	G/GT	3'- UTR (exon 6/7)	0.0261	-
U8	109,541,004	-	G/A	intronic	0.2023	-
	109,541,007	-	C/T	intronic	0.1603	-
	109,541,008	-	GA/AC GAC/GCC	intronic	0.6514	-
	109,541,009	-	A/C	intronic	0.1719	-
	109,541,043	-	A/G	intronic	0.0239	-
	109,541,051 - 109,541,052	-	AC/GT	intronic	0.1269	-
U10	109,541,435 ^109,541,436	rs144108564	C/CT C/TCT	intronic	0.1269	0.0976
	109,541,462 ^109,541,463	-	T/TA	intronic	0.1603	-

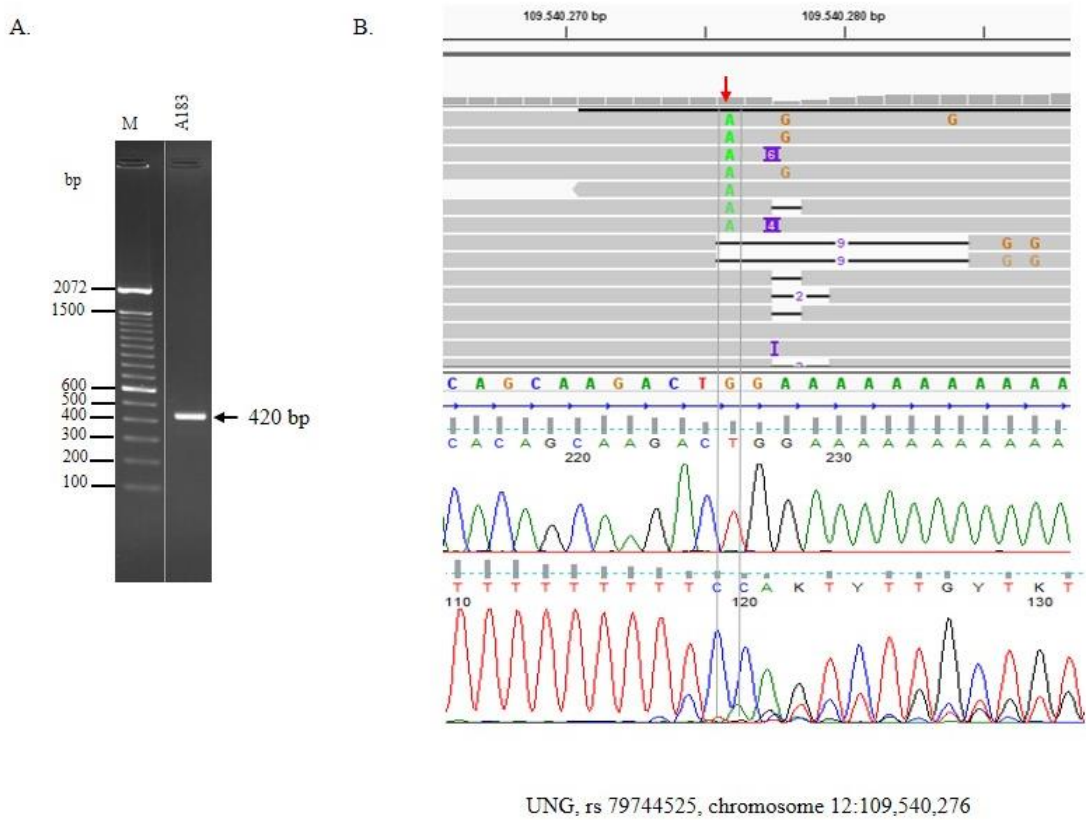


Figure 4.4: *UNG* gene, rs79744525, G/A variant confirmation by Sanger sequencing A. Agarose gel electrophoresis of 420 bp PCR products of *UNG* gene using U3 primer. Lane 1, "molecular size marker"; lane 6, 420 bp PCR product of A189 (sporadic AD patient number) B. A screenshot from the "IGV browser" showing G/A variation in reads for the sample A189 from "Ion-PGM sequencing experiment" (top panel). "Sanger sequencing chromatograms" of G/A variant showing false positive result; middle panel, Sanger sequencing using the forward U3 primer; lower panel, Sanger sequencing using the reverse U3 primer.

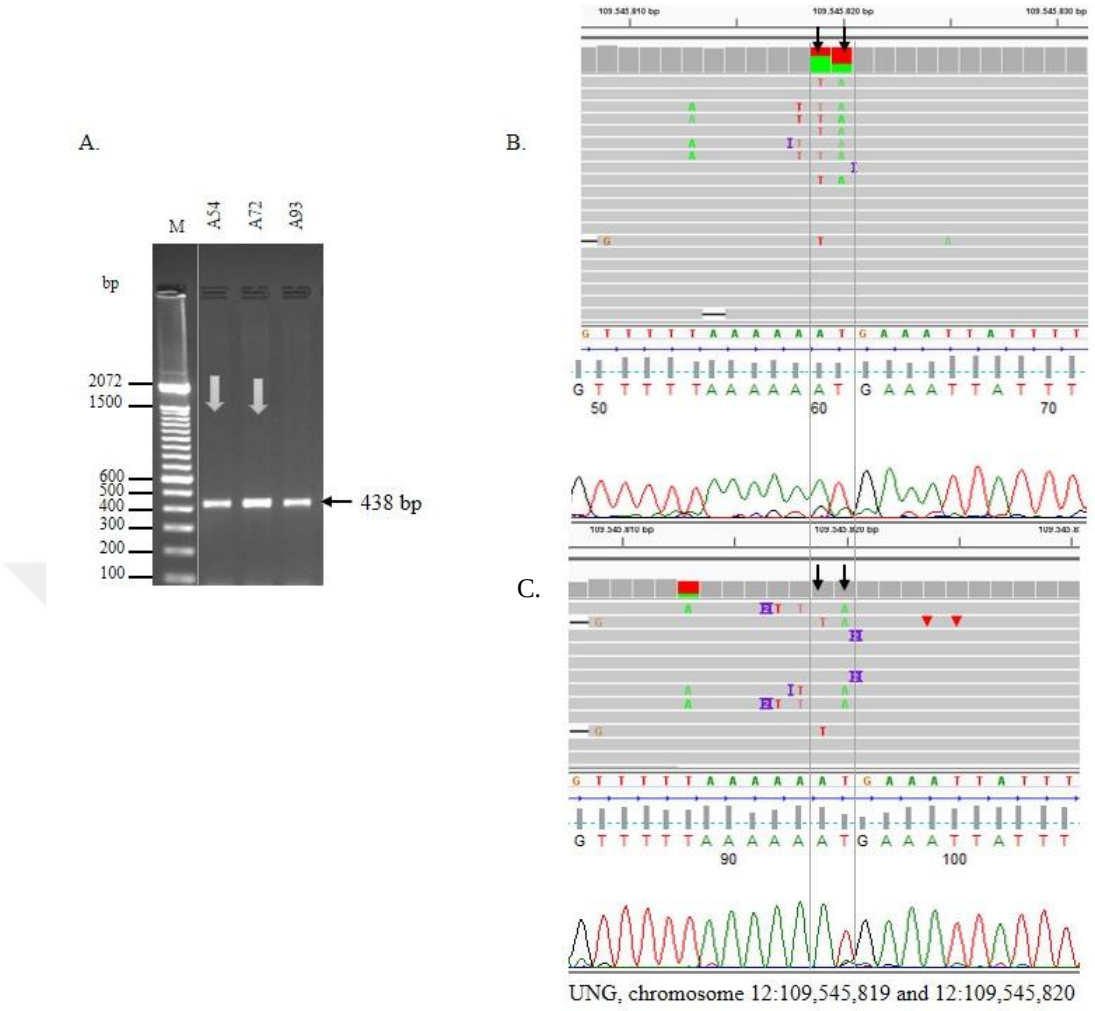


Figure 4.5: *UNG* gene, chromosome 12:109,545,819 and 109,545,820. AT/TA variant confirmation by Sanger sequencing A. "Agarose gel electrophoresis" of 438 bp PCR products of *UNG* gene using U5 primer. Lane 1, "molecular size marker"; lane 2 and 3, 420 bp PCR product of A54 and A72 (sporadic AD patient number) B. A screenshot from the "IGV browser" showing AT/TA variation in reads for the sample A54 from "Ion-PGM sequencing experiment" (top panel). "Sanger sequencing chromatograms" of AT/TA variant showing false positive result; lower panel "Sanger sequencing" using the forward U5 primer. C. A screenshot from the IGV browser showing AT/TA variation in reads for the sample A72 from "Ion-PGM sequencing experiment" (top panel). "Sanger sequencing chromatograms" of AT/TA variant showing false positive result; lower panel, "Sanger sequencing" using the forward U5 primer.

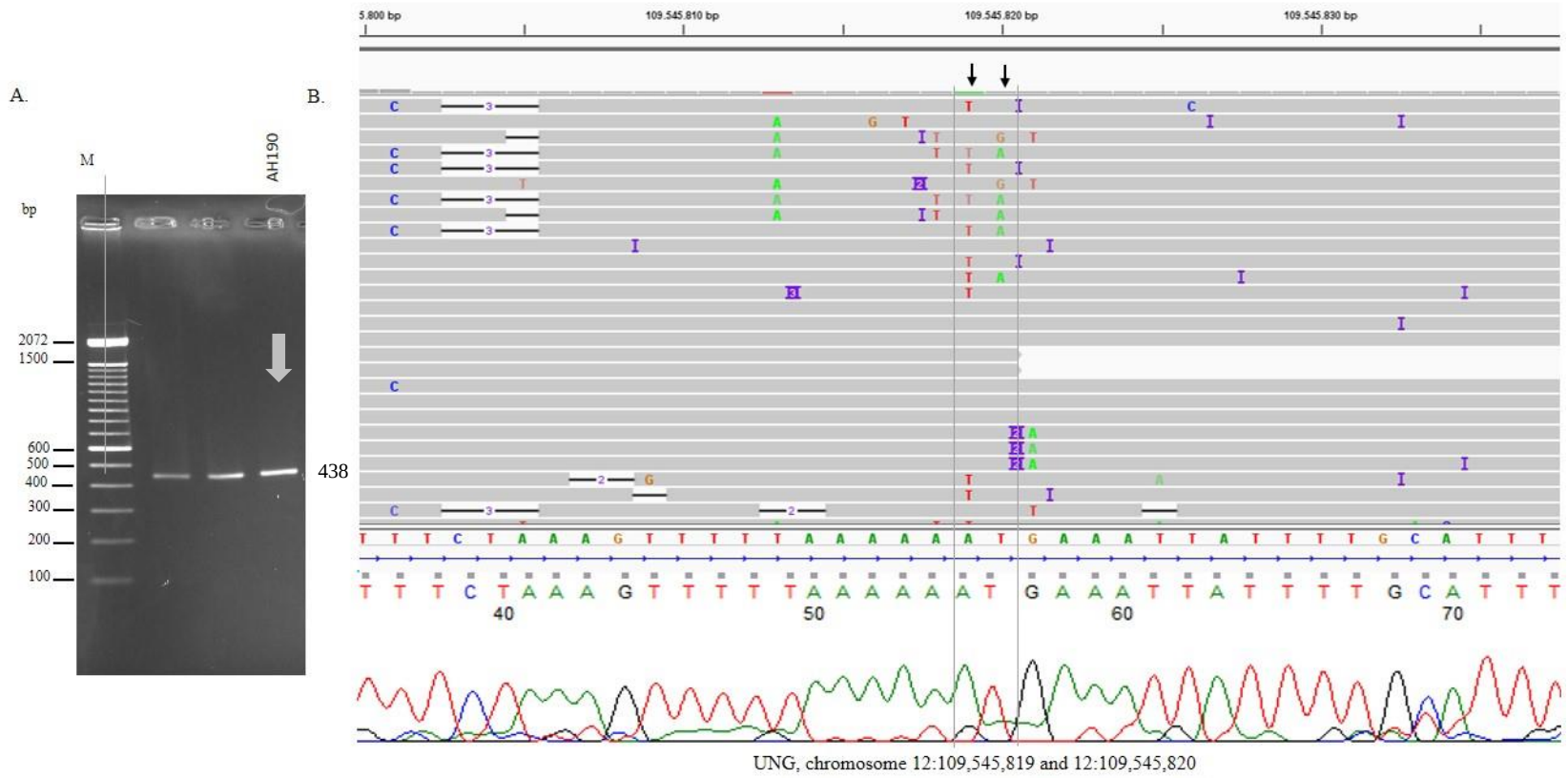


Figure 4.6: *UNG* gene, chromosome 12:109,545,819 and 12:109,545,820. AT/TA variant confirmation by "Sanger sequencing" A. "Agarose gel electrophoresis" of 438 bp PCR products of *UNG* gene using U5 primer. Lane 1, "molecular size marker"; lane 2 and lane 3, 420 bp PCR product of A190 (sporadic AD patient number) B. A screenshot from the IGV browser showing AT/TA variation in reads for the sample A190 from "Ion-PGM sequencing experiment" (top panel). "Sanger sequencing chromatograms" of AT/TA variant showing false positive result; lower panel, "Sanger sequencing" using the forward U5 primer.

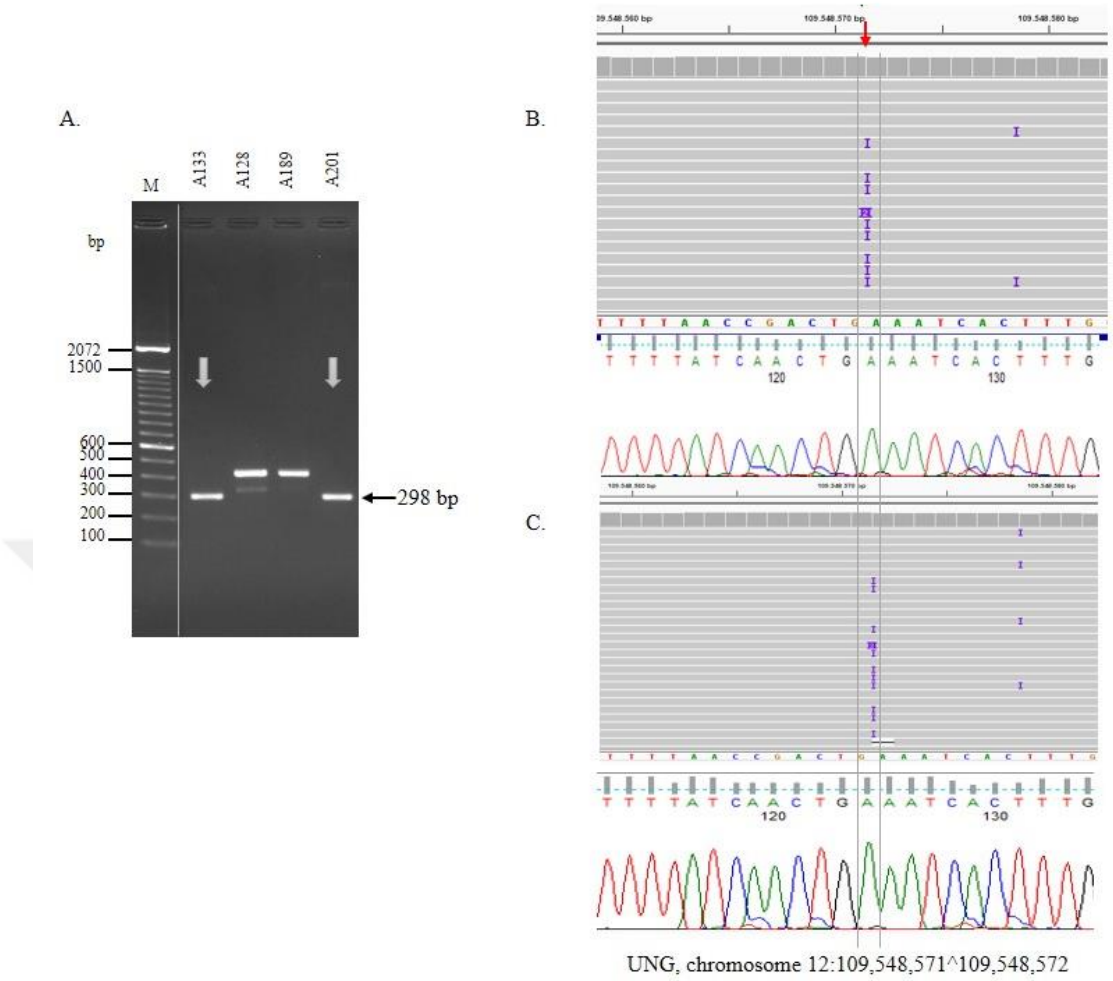


Figure 4.7: *UNG* gene, chromosome 12:109,548,571^109,548,572. Insertions' confirmation by "Sanger sequencing" A. "Agarose gel electrophoresis" of 298 bp PCR products of *UNG* gene using U7 primer. Lane 1, "molecular size marker"; lane 2 and lane 5, 420 bp PCR product of A133 and A201 (sporadic AD patient number) B. A screenshot from the "IGV browser" showing insertion in reads for the sample A133 from "Ion-PGM sequencing experiment" (top panel). "Sanger sequencing chromatograms" of insertion showing false positive result; lower panel, "Sanger sequencing" using the forward U7 primer. C. A screenshot from the "IGV browser" showing insertion in reads for the sample A201 from "Ion-PGM sequencing experiment" (top panel). "Sanger sequencing chromatograms" of insertion showing false positive result; lower panel, "Sanger sequencing" using the forward U7 primer. I: Insertion.

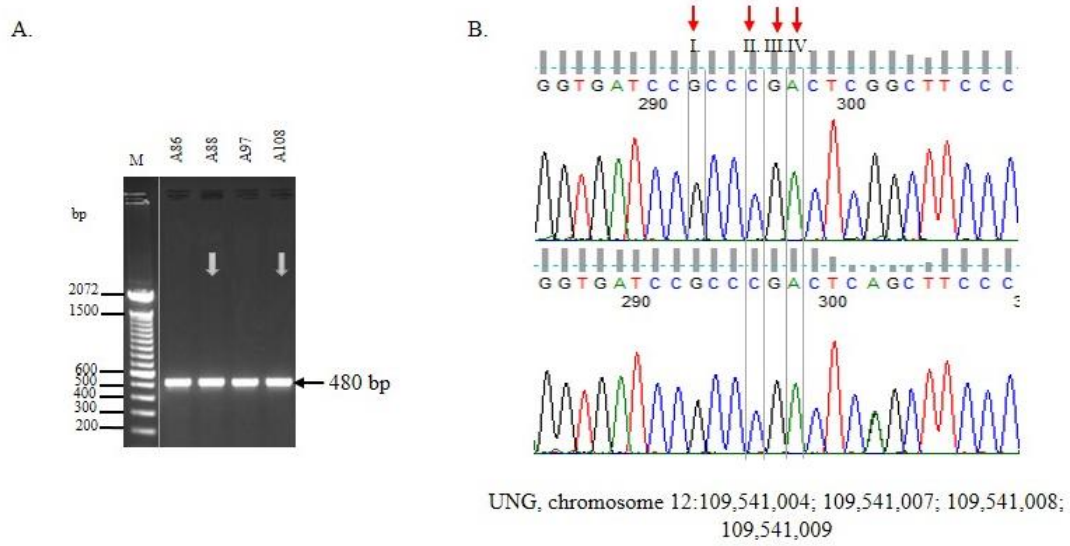


Figure 4.8: *UNG* gene, chromosome 12:109,541,004; 109,541,007; 109,541,009; 109,541,008. Variations' confirmation by "Sanger sequencing" A. "Agarose gel electrophoresis" of 480 bp PCR products of *UNG* gene using U8 primer. Lane 1, "molecular size marker"; lane 3 and lane 5, 480 bp PCR product of A88 and A108 (sporadic AD patient number) B. "Sanger sequencing chromatograms" of variations showing false positive result; top panel: A88, lower panel: A108. I, "Sanger sequencing" using the forward U8 primer. I, chromosome 12:109,541,004 G > A. II, chromosome 12:109,541,007 C > T. III, chromosome 12:109,541,008 GA > AC. IV, chromosome 12: 109,541,009 A > C.

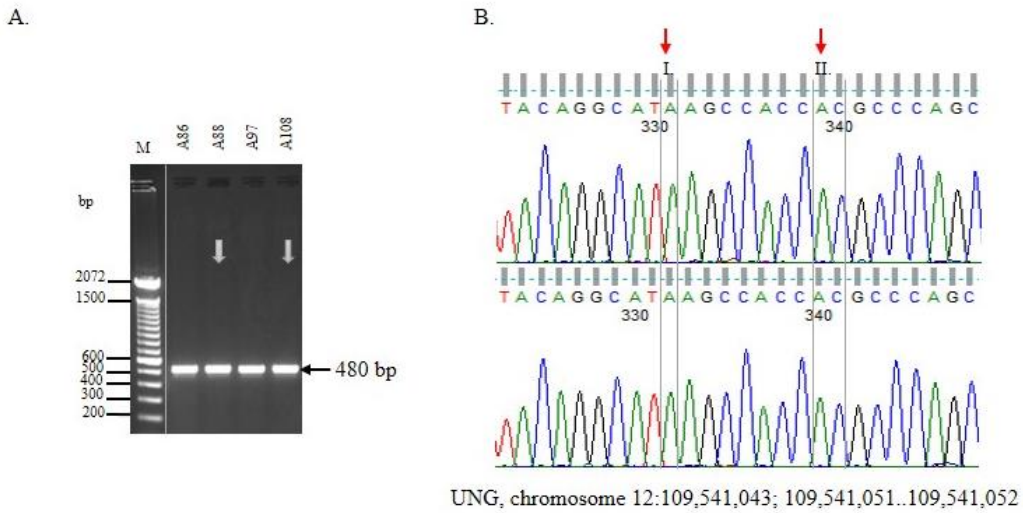


Figure 4.9: *UNG* gene, chromosome 12:109,541,043 and 109,541,051..109,541,052. Variations' confirmation by "Sanger sequencing" A. "Agarose gel electrophoresis" of 480 bp PCR products of *UNG* gene using U8 primer. Lane 1, "molecular size marker"; lane 3 and lane 5, 480 bp PCR product of A88 and A108 (sporadic AD patient number) B. "Sanger sequencing chromatograms" of variations showing false positive result; top panel: A88, bottom panel: A108. I, "Sanger sequencing" using the forward U8 primer. 12:109,541,043 A > G. II, chromosome 12:109,541,051..109,541,052 AC / GT.

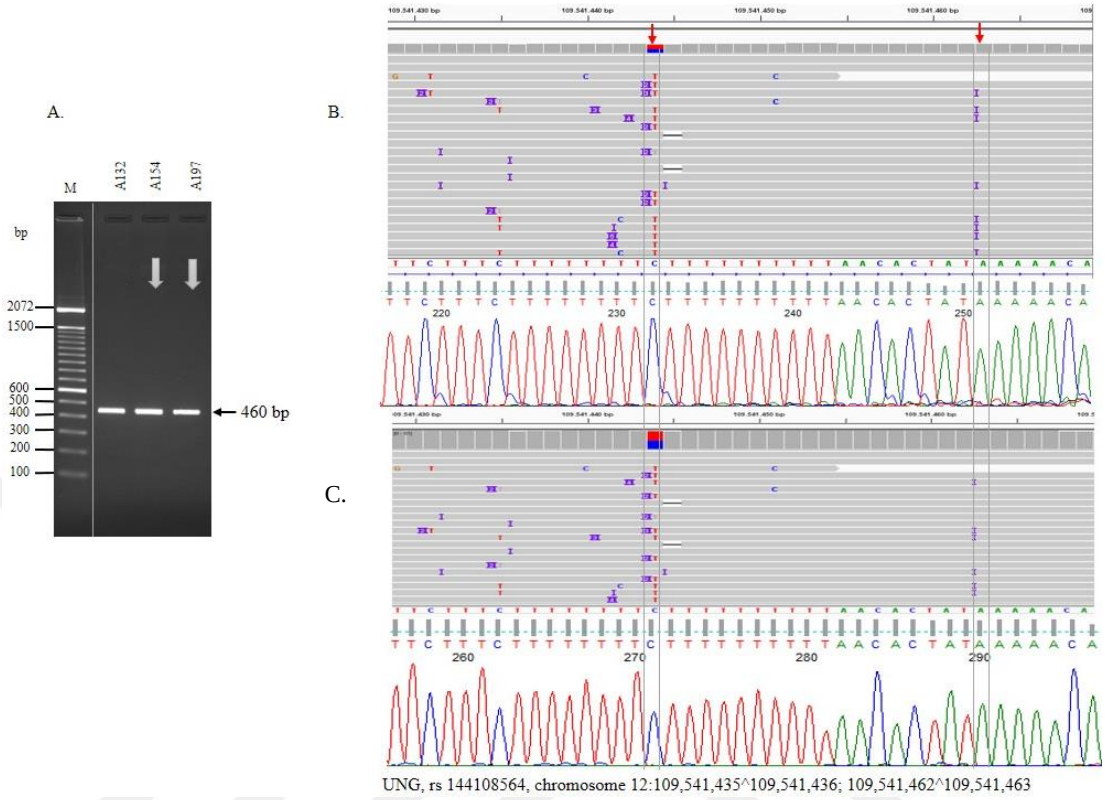


Figure 4.10: *UNG* gene, chromosome 12:109,541,435^109,541,436 and chromosome 12:109,541,462^109,541,463, insertions confirmation by "Sanger sequencing" A. "Agarose gel electrophoresis" of 460 bp PCR products of *UNG* gene using U10 primer. Lane 1, "molecular size marker"; lane 3 and lane 4, 460 bp PCR product of A154 and A197 (sporadic AD patient number) B. A screenshot from the "IGV browser" showing variations in reads for the sample A154 from Ion-PGM sequencing experiment (top panel). "Sanger sequencing chromatograms" of insertion variants showing false positive result; lower panel, "Sanger sequencing" using the forward U10 primer. C. A screenshot from the "IGV browser" showing variations in reads for the sample A197 from "Ion-PGM sequencing experiment" (top panel). "Sanger sequencing chromatograms" of insertion variants showing false positive result; lower panel, "Sanger sequencing" using the forward U10 primer.

Table 4.4: The positive results.

Primer's ID	Location, chromosome12	rs number	Genotype	Intron / exon	Fisher exact test p-value	MAF
U2-V2	109,535,809	rs1018782	A/G	Intronic	0.0948	0.20627
	109,535,857	rs1018783	T/A	Intronic	0.2081	0.194688
U4	109,537,423	rs1610925	T/TTA	Intronic	0.0201	0.232
U5	109,545,951	rs2430678	A/G	Intronic	0.2023	0.045927
U5.1	109,546,114	rs2515913	C/T	Intronic	0.2023	0.040136
U6	109,546,886	rs2268406	T/G	Intronic	0.0491	0.142971

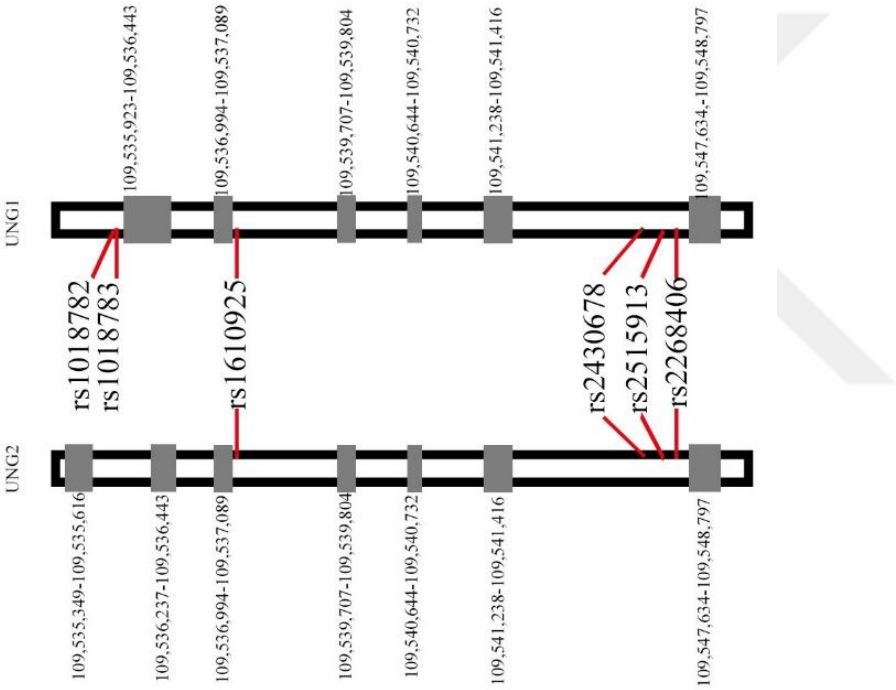


Figure 4.11: The *UNG* gene schema showing the verified variants.

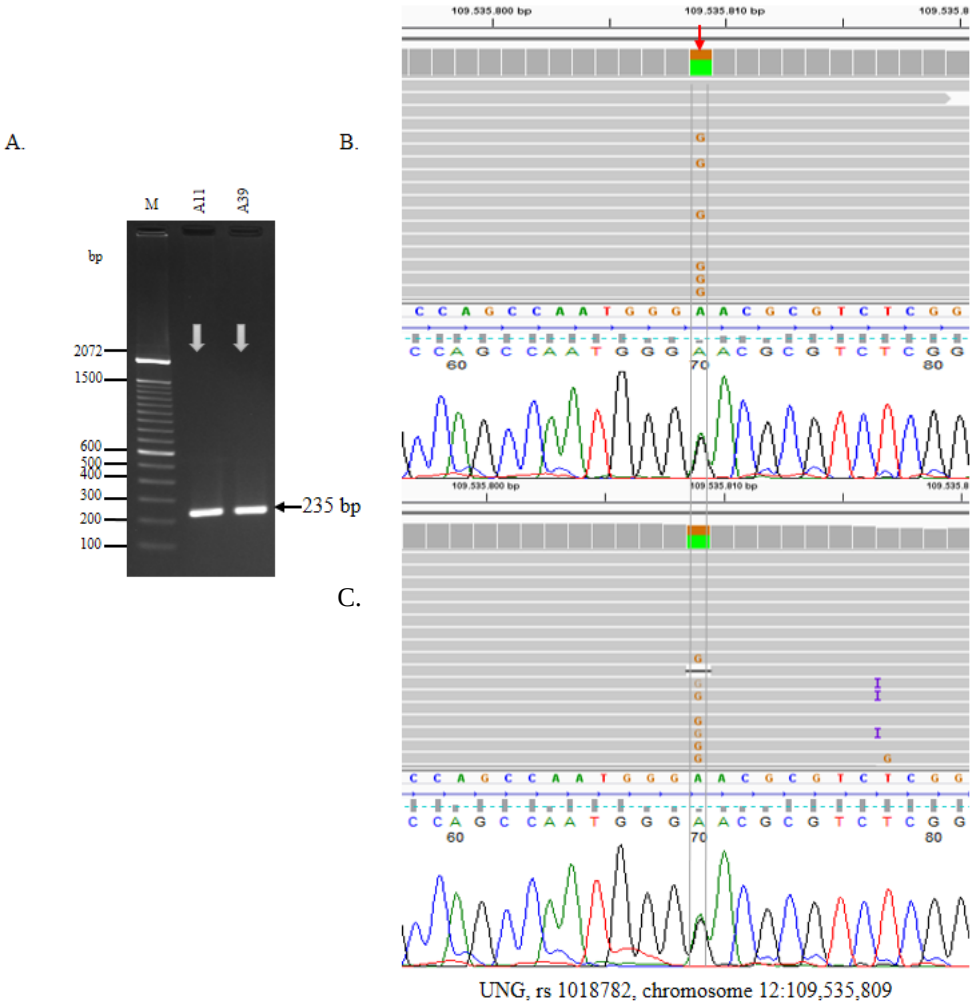


Figure 4.12: *UNG* gene, rs 1018782, G/A variant confirmation by "Sanger sequencing" A. "Agarose gel electrophoresis" of 235 bp PCR products of *UNG* gene using U2-V2 primer. Lane 1, "molecular size marker"; lane 2 and lane 3, 235 bp PCR product of A11 and A39 (sporadic AD patient number) B. A screenshot from the "IGV browser" showing G/A variation in reads for the sample A11 from "Ion-PGM sequencing experiment" (top panel). "Sanger sequencing chromatograms" of G/A variant showing positive result; middle panel, "Sanger sequencing" using the forward U2-V2 primer; lower panel, "Sanger sequencing" using the forward U2-V2 primer. C. A screenshot from the "IGV browser" showing G/A variation in reads for the sample A39 from "Ion-PGM sequencing experiment" (top panel). Sanger sequencing chromatograms of G/A variant showing positive result; lower panel, "Sanger sequencing" using the forward U2-V2 primer.

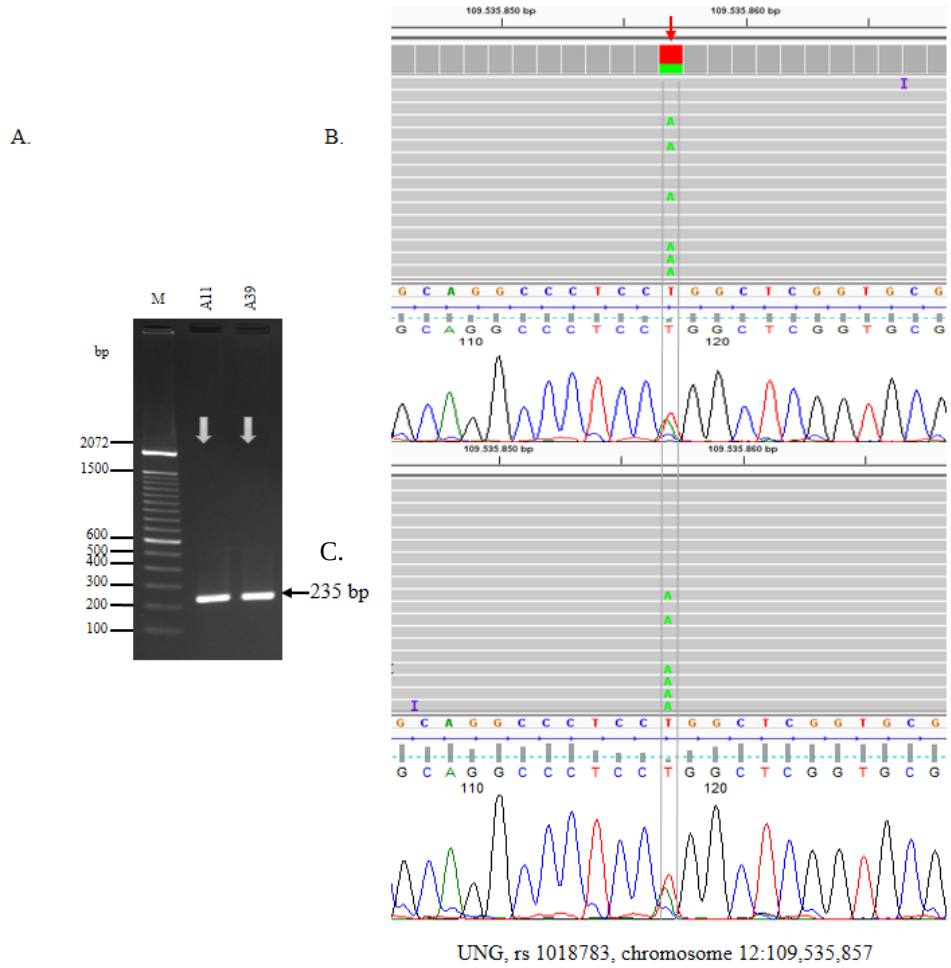


Figure 4.13: *UNG* gene, rs 1018783, T/A variant confirmation by "Sanger sequencing" A. "Agarose gel electrophoresis" of 235 bp PCR products of *UNG* gene using U2-V2 primer. Lane 1, "molecular size marker"; lane 2 and lane 3, 235 bp PCR product of A11 and A39 (sporadic AD patient number) B. A screenshot from the "IGV browser" showing G/A variation in reads for the sample A11 from "Ion-PGM sequencing experiment" (top panel). "Sanger sequencing chromatograms" of T/A variant showing positive result; middle panel, "Sanger sequencing" using the forward U2-V2 primer; lower panel, "Sanger sequencing" using the forward U2-V2 primer. C. A screenshot from the "IGV browser" showing T/A variation in reads for the sample A39 from "Ion-PGM sequencing experiment" (top panel). "Sanger sequencing chromatograms" of T/A variant showing positive result; middle panel, "Sanger sequencing" using the forward U2-V2 primer; lower panel, "Sanger sequencing" using the forward U2-V2 primer.

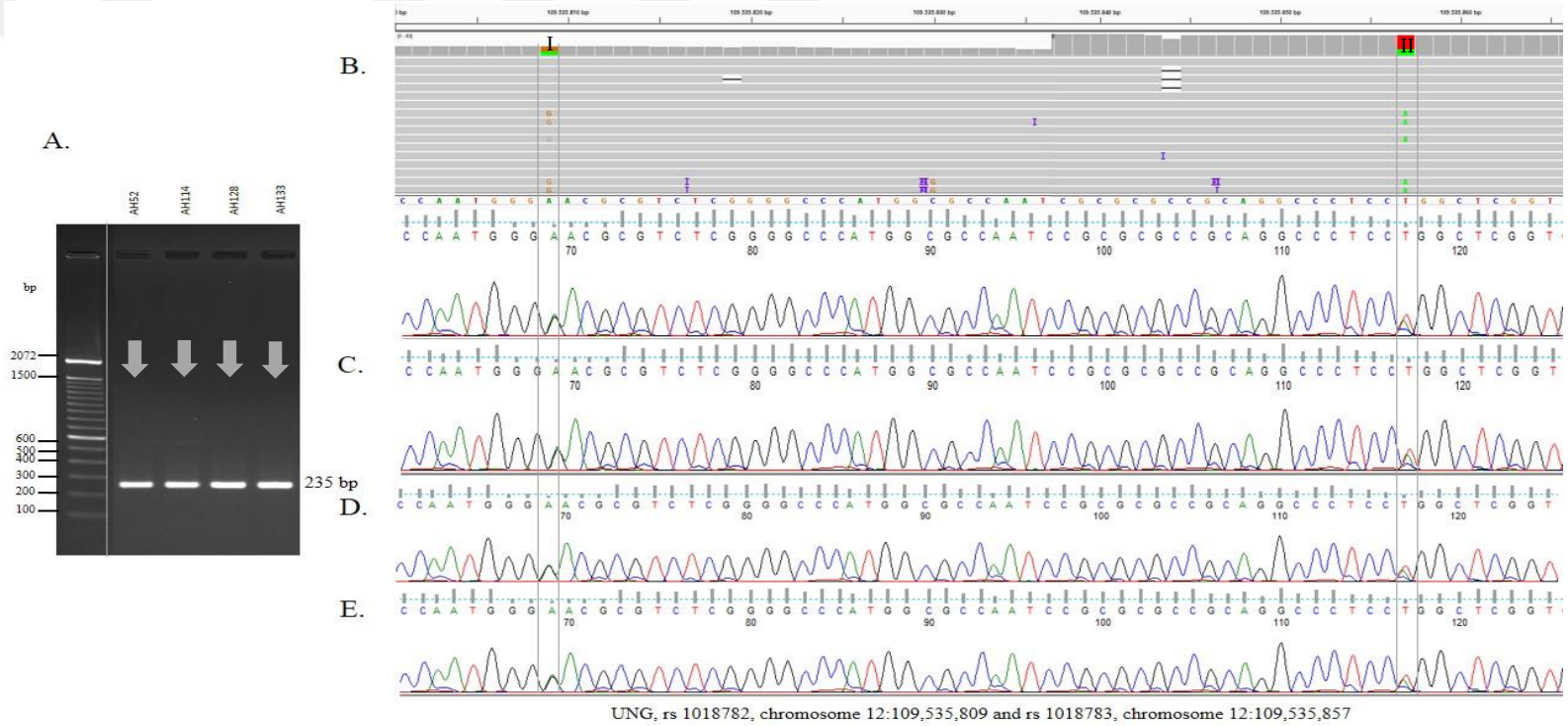


Figure 4.14: *UNG* gene, rs 1018783, G/A variant and rs 1018783, T/A variant confirmation by "Sanger sequencing" A. "Agarose gel electrophoresis" of 235 bp PCR products of *UNG* gene using U2-V2 primer. Lane 1, "molecular size marker"; lane 2, 3, 4, and 5, 235 bp PCR product of A52, A114, A128 and A133 (sporadic AD patient number) B. A screenshot from the "IGV browser" showing G/A and T/A variation in reads for the sample A52 from "Ion-PGM sequencing experiment" (top panel). "Sanger sequencing chromatograms" of variants showing positive result; lower panel, "Sanger sequencing" using the forward U2-V2 primer. C. "Sanger sequencing chromatogram" of variants showing positive result for A114, Sanger sequencing using the forward U2-V2 primer. D. "Sanger sequencing chromatograms" of variants showing positive result for A128; "Sanger sequencing" using the forward U2-V2 primer. E." Sanger sequencing chromatograms" of variants showing positive result for A133; "Sanger sequencing" using the forward U2-V2 primer. I: rs 1018782, G/A, II: rs1018783, T/A.

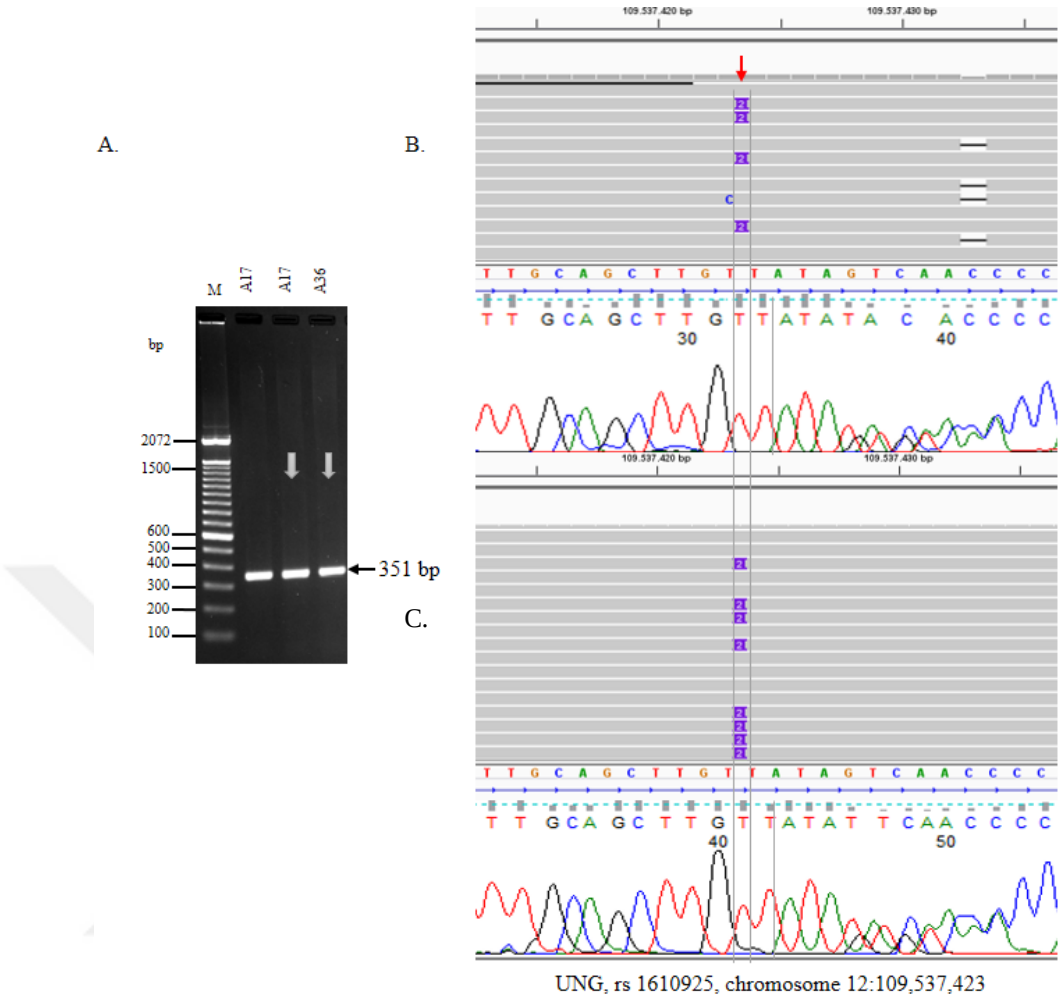


Figure 4.15: *UNG* gene, rs 1610925, TA insertion confirmation by "Sanger sequencing" A. "Agarose gel electrophoresis" of 351 bp PCR products of *UNG* gene using U4 primer. Lane 1, "molecular size marker"; lane 3 and 4, 351 bp PCR product of A17 and A36 (sporadic AD patient number) B. A screenshot from the "IGV browser" showing TA insertion in reads for the sample A17 from "Ion-PGM sequencing experiment" (top panel). "Sanger sequencing chromatograms" of TA insertion showing positive result; lower panel, "Sanger sequencing" using the forward U4 primer. C. A screenshot from the "IGV browser" showing TA insertion in reads for the sample A36 from "Ion-PGM sequencing experiment" (top panel). "Sanger sequencing chromatograms" of TA insertion variant showing positive result; lower panel, "Sanger sequencing" using the forward U4 primer.

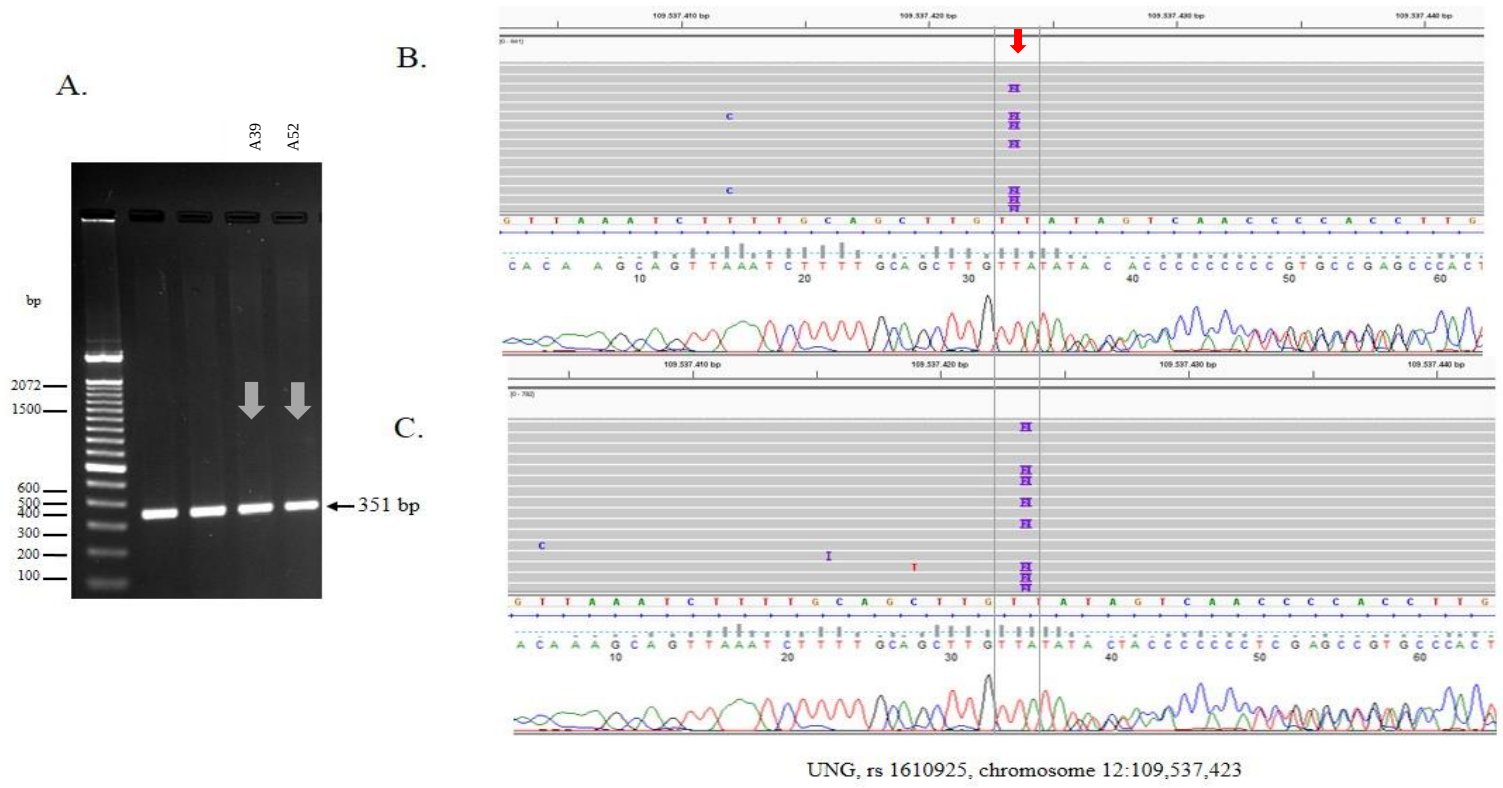


Figure 4.16: *UNG* gene, rs1610925, TA insertion confirmation by "Sanger sequencing" A. "Agarose gel electrophoresis" of 351 bp PCR products of *UNG* gene using U4 primer. Lane 1, "molecular size marker"; lane 4 and 5, 351 bp PCR product of A39 and A52 (sporadic AD patient number) B. A screenshot from the "IGV browser" showing TA insertion in reads for the sample A39 from "Ion-PGM sequencing experiment" (top panel). Sanger sequencing chromatograms of TA insertion showing positive result; lower panel, "Sanger sequencing" using the forward U4 primer. C. A screenshot from the IGV browser showing TA insertion variation in reads for the sample A52 from "Ion-PGM sequencing experiment" (top panel). "Sanger sequencing chromatogram" of TA insertion variant showing positive result; lower panel, Sanger sequencing using the forward U4 primer.

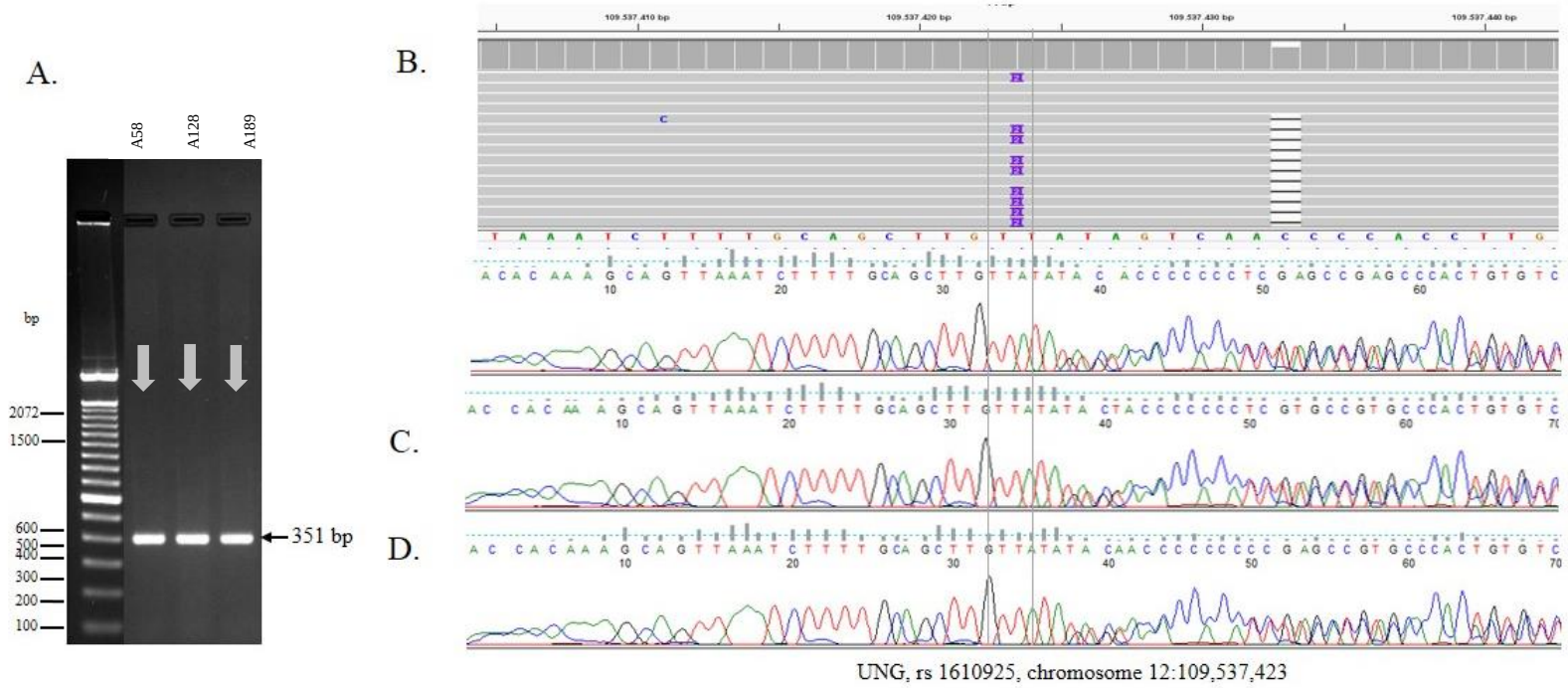


Figure 4.17: *UNG* gene, rs1610925, TA insertion confirmation by "Sanger sequencing" A. "Agarose gel electrophoresis" of 351 bp PCR products of *UNG* gene using U4 primer. Lane 1, "molecular size marker"; lane 4 and 5, 351 bp PCR product of A58, A128 and A189 (sporadic AD patient number) B. A screenshot from the "IGV browser" showing TA insertion in reads for the sample A58 from "Ion-PGM sequencing experiment" (top panel). "Sanger sequencing chromatograms" of TA insertion showing positive result; lower panel, "Sanger sequencing" using the forward U4 primer. C. "Sanger sequencing chromatogram" of TA insertion variant showing positive result for sample A128; "Sanger sequencing" using the forward U4 primer. D. "Sanger sequencing chromatogram" of TA insertion variant showing positive result for sample A189; "Sanger sequencing" using the forward U4 primer.

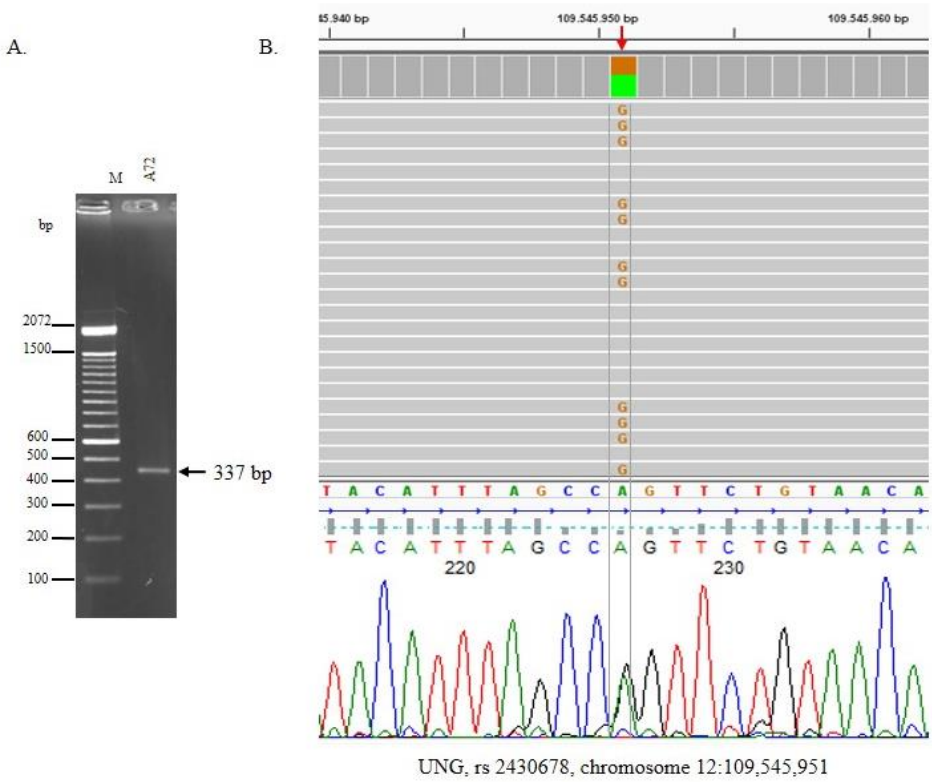


Figure 4.18: *UNG* gene, rs 2430678, A/G variation confirmation by "Sanger sequencing" A. "Agarose gel electrophoresis" of 337 bp PCR products of *UNG* gene using U5 primer. Lane 1, "molecular size marker"; lane 2, 337 bp PCR product of A72 (sporadic AD patient number) B. A screenshot from the "IGV browser" showing G/A variation in reads for the sample A72 from "Ion-PGM sequencing experiment" (top panel). "Sanger sequencing chromatogram" of A/G variation showing positive result; lower panel, "Sanger sequencing" using the forward U5 primer.

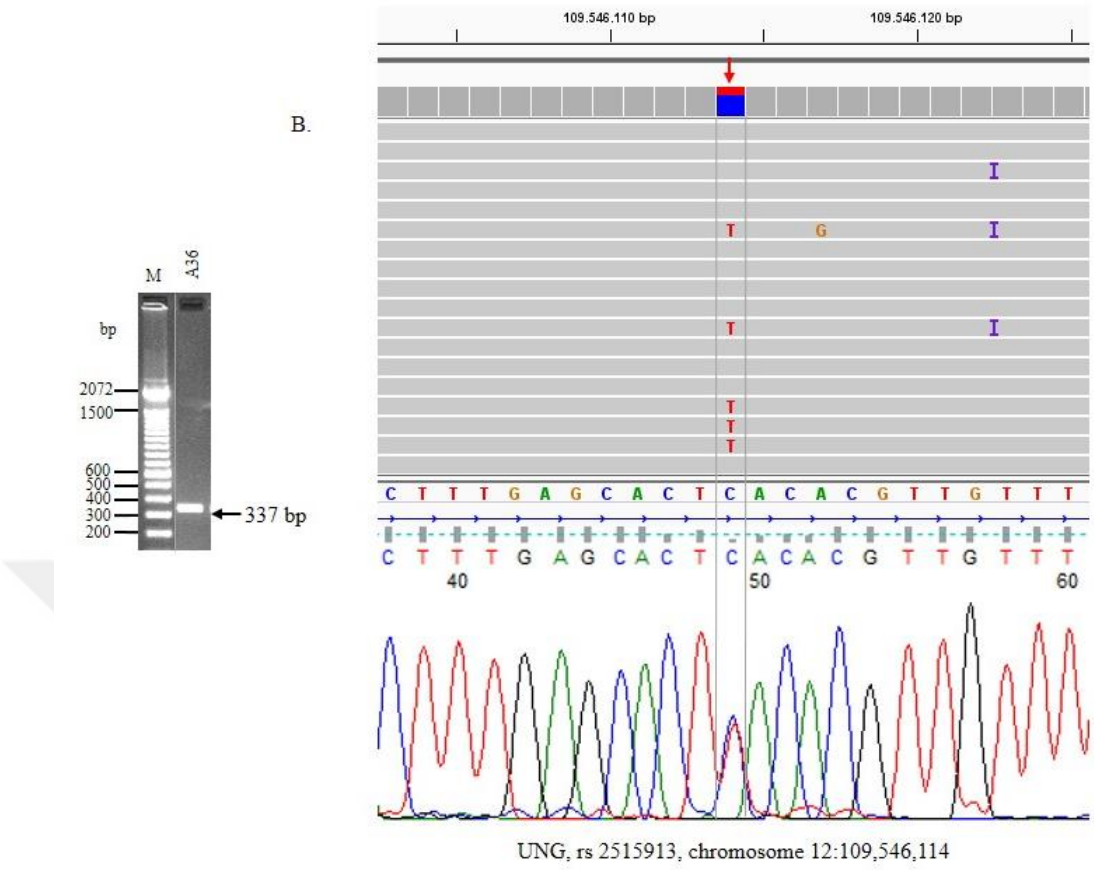


Figure 4.19: *UNG* gene, rs2515913, C / T variation confirmation by "Sanger sequencing" A. "Agarose gel electrophoresis" of 337 bp PCR products of *UNG* gene using U5.1 primer. Lane 1, "molecular size marker"; lane 2, 337 bp PCR product of A36 (sporadic AD patient number) B. A screenshot from the IGV browser showing C / T variation in reads for the sample A36 from "Ion-PGM sequencing experiment" (top panel). "Sanger sequencing chromatogram" of C / T variation showing positive result; lower panel, "Sanger sequencing" using the forward U5.1 primer.

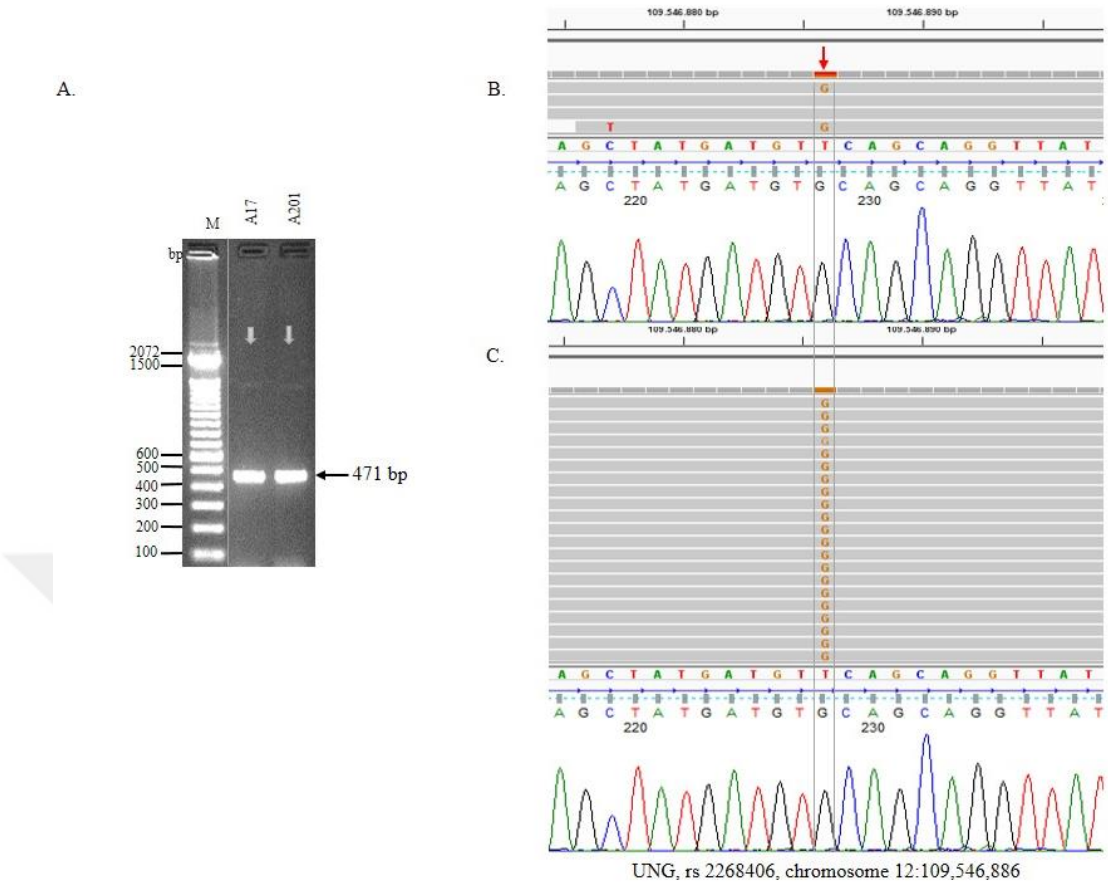


Figure 4.20: *UNG* gene, rs 2268406, T > G variation confirmation by "Sanger sequencing" A. "Agarose gel electrophoresis" of 471 bp PCR products of *UNG* gene using U6 primer. Lane 1, "molecular size marker"; lane 2 and 3, 471 bp PCR product of A17 and A201 (sporadic AD patient number) B. A screenshot from the "IGV browser showing" T > G variation in reads for the sample A17 from "Ion-PGM sequencing experiment" (top panel). "Sanger sequencing chromatogram" of T>G variation showing positive result; lower panel, "Sanger sequencing" using the forward U6 primer. C. A screenshot from the "IGV browser" showing T > G variation in reads for the sample A201 from "Ion-PGM sequencing experiment" (top panel). "Sanger sequencing chromatogram" of T>G variation showing positive result; lower panel, "Sanger sequencing" using the forward U6 primer.

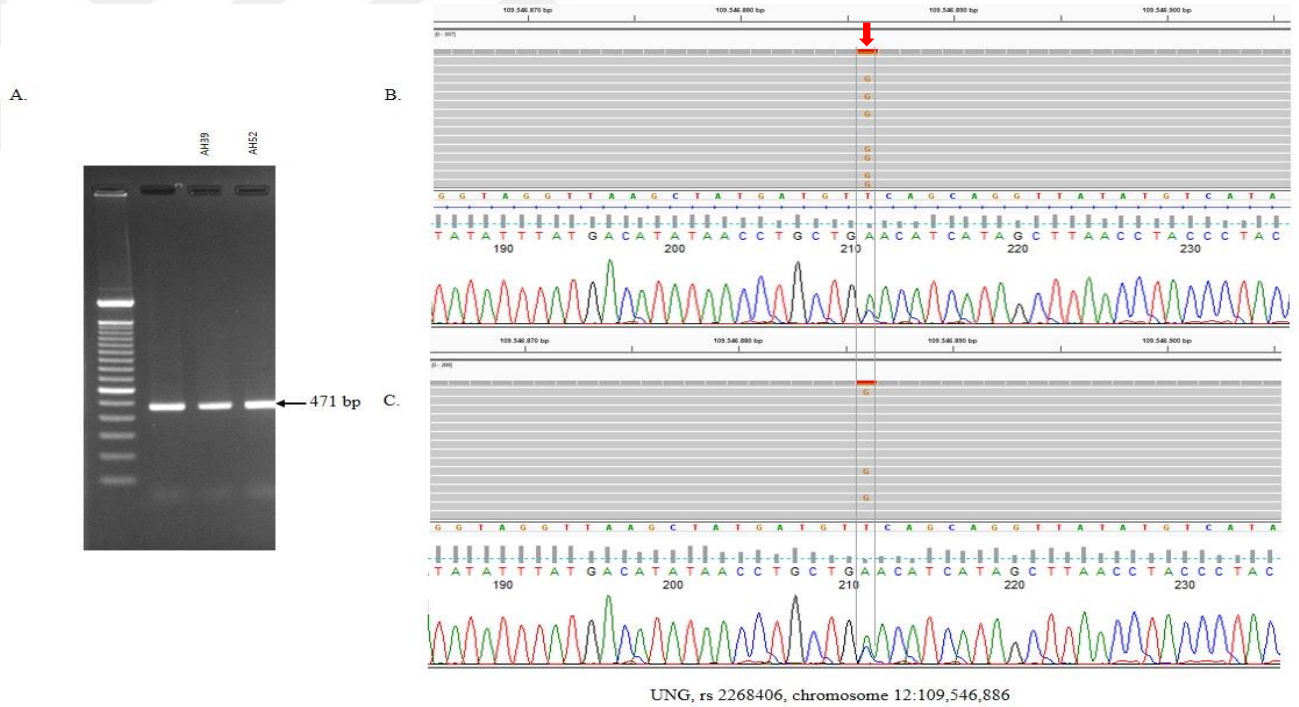


Figure 4.21: *UNG* gene, rs2268406, T>G variation confirmation by "Sanger sequencing" A. "Agarose gel electrophoresis" of 471 bp PCR products of *UNG* gene using U6 primer. Lane 1, "molecular size marker"; lane 3 and 4, 471 bp PCR product of A39 and A52 (sporadic AD patient number) B. A screenshot from the "IGV browser" showing T>G variation in reads for the sample A39 from "Ion-PGM sequencing experiment" (top panel). "Sanger sequencing chromatogram" of T>G variation showing positive result; lower panel, Sanger sequencing using the reverse U6 primer. C. A screenshot from the IGV browser showing T > G variation in reads for the sample A52 from "Ion-PGM sequencing experiment" (top panel). "Sanger sequencing chromatogram" of T > G variation showing positive result; lower panel, "Sanger sequencing" using the reverse U6 primer.

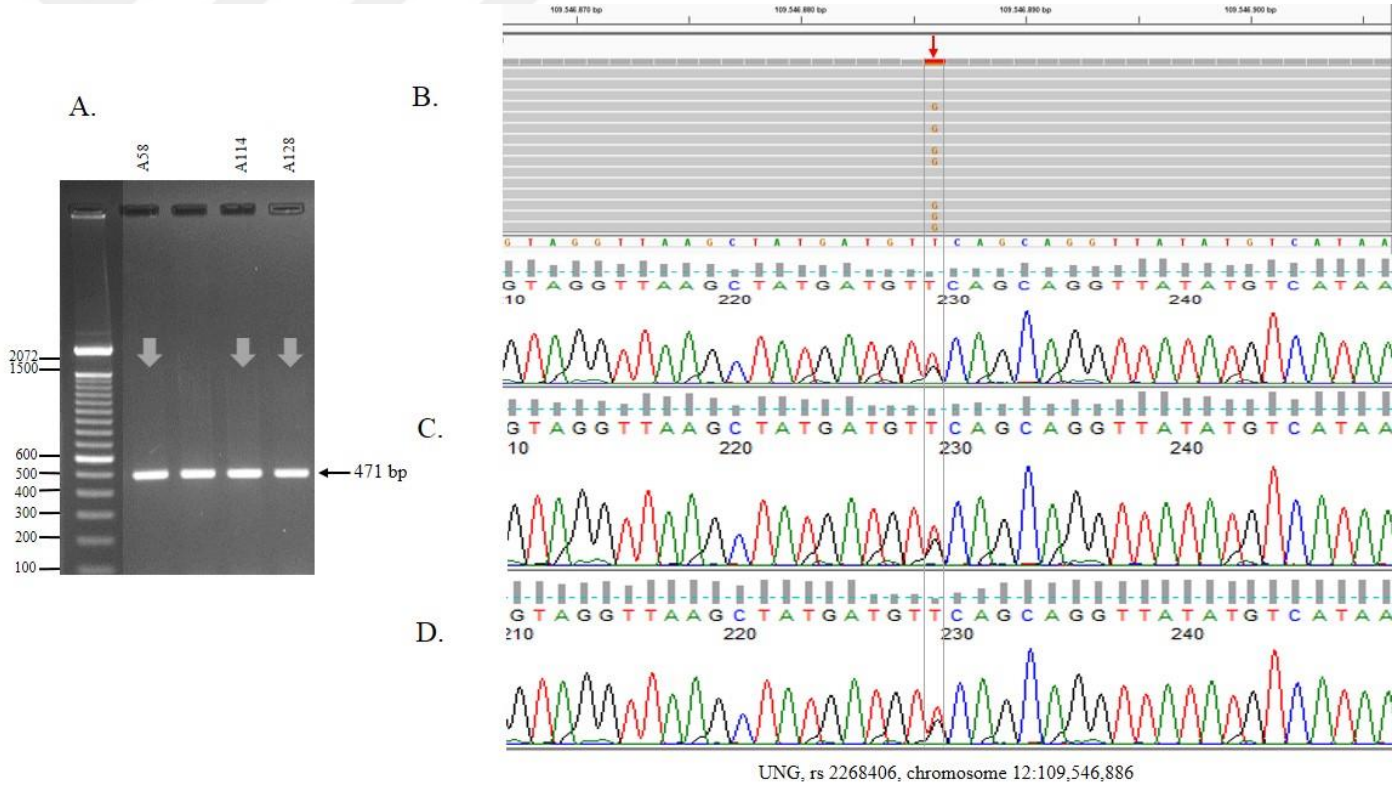


Figure 22: *UNG* gene, rs 2268406, T > G variation confirmation by "Sanger sequencing" A. "Agarose gel electrophoresis" of 471 bp PCR products of *UNG* gene using U6 primer. Lane 1, "molecular size marker"; lane 2, 4 and 5, 471 bp PCR product of A58, A114 and A128 (sporadic AD patient number) B. A screenshot from the "IGV browser" showing T > G variation in reads for the sample A58 from "Ion-PGM sequencing experiment" (top panel). "Sanger sequencing chromatogram" of T>G variation showing positive result; lower panel, "Sanger sequencing" using the forward U6 primer. C. "Sanger sequencing chromatogram" of T > G variation showing positive result for A114; lower panel, "Sanger sequencing" using the forward U6 primer. D. "Sanger sequencing chromatogram" of T > G variation showing positive result for A128; lower panel, "Sanger sequencing" using the forward U6 primer.

In this study, “Hardy–Weinberg Equilibrium (HWE)” was used for the calculation of significance of the variants. Observed and expected allele frequencies were determined with this equilibrium for patients and controls (Table 4.5). Pearson χ^2 and p values were calculated (93). The Odds ratio was calculated with the confidence intervals (CI) of 95% confidence and a p value <0.05 was considered statistically significant (94,95) (Table 4.5). Fisher exact test p value is two-way.

4.4.Epistatic Relations

In this study, observed variations compared for epistatic relations. The binary and multiple epistatic relations are presented in Table 4.6 and Table 4.7.

4.5. Common Significant Variants Independent from Sporadic AD Outcome

Independent from sporadic AD outcome, there are common *UNG* variants were identified in all 248 samples (Table 4.8 and Fıgure 23). For example, rs246078 were determined in 97.6% of all the samples (94.78%-G/G homozygous and 2.82% A/G heterozygous SNP). Rs2541886 were determined in 70.97% of all the samples (28.63%-T/T homozygous and 42.34% G/T heterozygous SNP).

Table 4.5. Allele and genotype frequency distribution of 6 *UNG* variants in Sporadic AD patients and healthy controls.

	Allele frequency					Genotype frequency				
UNG variant	**Allele**	**Case**	**Control**	**P value**	**OR (95 % CI)**	**Genotype**	**Case**	**Control**	**P value**	**Pearson χ^2**
rs1018782	A	375	206	0.474 (0.899)	1.051 (0.640-1.726)	A/A	158	87	0.841	0.04
	G	67	34			A/G	59	32		
						G/G	4	1		
rs1018783	T	380	206	0.715 (0.699)	0.895 (0.541-1.479)	T/T	165	87	0.663	0.19
	A	62	34			T/A	50	32		
						A/A	6	1		
rs1610925	T	323	91	0.020 (0.039)	2.330 (1.069-5.079)	T/T	131	41	0.030	4.712
	TTA	73	9			T/TTA	61	9		
						TTA/TTA	6	0		
rs2430678	A	389	100	(0.350)	-	A/A	191	50	-	-
	G	7	0			A/G	7	0		
						G/G	0	0		
rs2515913	C	389	100	(0.350)	-	C/C	191	50	-	-
	T	7	0			C/T	7	0		
						T/T	0	0		
rs2268406	T	331	91	0.049 (0.080)	2.028 (0.928- 4.434)	T/T	137	41	0.072	3.232
	G	65	9			T/G	57	9		
						G/G	4	0		

Table 4.6: Binary epistasis relations.

Variation 1		**Variation 2**		**Epistasis relation**		
Rs number	AD / Control	Rs number	AD / Control	AD / Control	All affected patients (%)	All affected controls (%)
rs1610925	67/9	rs2268406	61/9	55/9	27.7%	18%
		rs1018782	56/9	50/9	25.3%	18%
		rs1018783	56/9	43/9	21.7%	18%
rs2268406	61/9	rs1018782	56/9	52/9	26.3%	18%
		rs1018783	56/9	45/9	22%	18%
rs1018782	63/33	rs1018783	56/33	52/33	29%	28%

Table 4.7: Multiple epistasis relations.

Variation 1		**Variation 2**		**Variation 3**		**Variation 4**		**Variation 5**		**Epistasis relation**		
Rs number	AD / Control	Rs number	AD / Control	Rs number	AD / Control	Rs number	AD / Control	Rs number	AD / Control	AD / Control	All affected patients (%)	All affected controls (%)
rs1018782	56/9	rs1018783	56/9	rs1610925	67/9	rs2268406	61/9	-	-	49/9	22.7%	18%
rs1018782	56/9	rs1018783	56/9	rs1610925	67/9	rs2268406	61/9	rs80001089	42/5	35/5	17.6%	10%

Table 4.8. Common significant variants independent from sporadic AD outcome.

UNG gene, chromosome 12	Type	dbSNP (rs)	Ref	Allele	Genotype	Intron/ exon	Zygosity	AD count	Control Count	Total samples	Percentage of the variants in the sample group	MAF
109,543,714	SNP	rs246078	A	G	A/G	intronic	Heterozygous	4/198	3/50	7/248	2.82	0,004792
					G/G		Homozygous	193/198	42/50	235/248	94.76	
109,537,564	SNP	rs2541886	G	T	G/T	intronic	Heterozygous	82/198	23/50	105/248	42.34	0.247604
					T/T		Homozygous	56/198	15/50	71/248	28.63	
109,547,060	SNP	rs246079	A	G	A/G	intronic	Heterozygous	90/198	19/50	109/248	43.95	0.383586
					G/G		Homozygous	29/198	6/50	35/248	14.11	
109,542,392	SNP	rs3219243	T	C	T/C	intronic	Heterozygous	53/198	14/50	67/248	27.02	0.228235
					C/C		Homozygous	4/198	2/50	6/248	2.42	
109,536,559	SNP	rs3219211	A	C	A/C	intronic	Heterozygous	52/198	14/50	66/248	26.61	0.226637
					C/C		Homozygous	4/198	2/50	6/248	2.42	
109,540,613^ 109,540,614	Insertion	rs3219235	-	T	G/GT	intronic	Heterozygous	37/198	9/50	43/248	17.34	-
					GT/GT		Homozygous	2/198	2/50	4/248	1.61	
					GT/GTT		Heterozygous	2/198	0/50	2/248	0.8	
109,541,171	SNP	rs2569987	T	C	T/C	intronic	Heterozygous	29/198	9/50	38/248	15.32	0.059505
					C/C		Homozygous	1/198	0/50	1/248	0.4	0.059505
109,548,571^ 109,548,572	Insertion		G	GT	G/GT	3'-UTR (exon 7)	Heterozygous	23/198	1/50	24/248	9.67	-

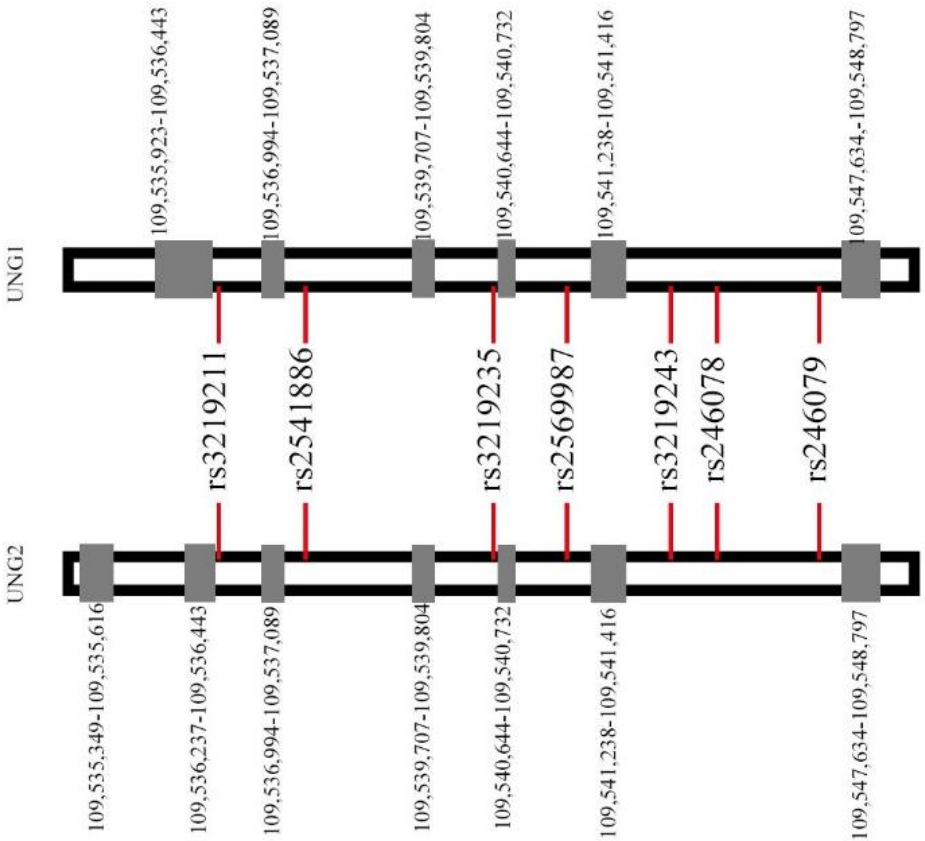


Figure 4.23. The location of common variants on *UNG* gene structure.

5. DISCUSSION and CONCLUSION

Aging and environmental factors play a role in the etiology of sporadic AD. Since it has been shown that early diagnosis slow the progression of sporadic AD and reduce its symptoms, screening studies for the identification of genetic risk factors for sporadic AD has increased. So far, APOE variants are known to be a genetic risk factor for sporadic AD however, since these variants do not show very strong association with sporadic AD, they have not been used widely in the diagnosis of the disease. The oxidative DNA damage accumulation and the defect in BER mechanism for repairing oxidative DNA damage especially in nerve cells plays an important role in the etiology and pathogenesis of various neurodegenerative diseases, including sporadic AD. Thus, in recent years, BER genes have become the focus of research to determine genetic risk factors for sporadic AD, but BER gene/gene variant(s) related to sporadic AD has not been determined yet. The biochemical, cellular, molecular and behavioral studies are done with the post-mortem brain tissues of sporadic Alzheimer patients, Alzheimer mouse models and cell lines, a specific BER gene knockout mice models and cell lines revealed the strong relation between BER deficiency and sporadic AD pathogenesis (6–9).

UDG is a monofunctional DNA glycosylase involved in the first step of both the nuclear and mitochondrial BER pathway. *UNG* gene encodes both nuclear and mitochondrial isoform of UDG formed by alternative splicing (4,5). The lack of UDG protein due to *UNG* gene silencing in rat hippocampal neurons caused neuronal death by inducing neuronal apoptosis, suggesting that this protein plays a crucial role in the neuronal development (6). It has been shown that folic acid deficiency and high homocysteine levels (7), which increase the risk of AD, enhance DNA damage in nerve cells, in particular uracil misincorporation and hypomethylation of DNA, and thus make neurons susceptible to amyloid AβPP toxicity (8). The studies in Alzheimer's mouse models and cell lines have shown that the damage to the BER pathway due to the decrease in uracil incision activity of the UDG protein also contributes to the increase in neuronal hippocampal neurodegeneration and neuronal AβPP toxicity (6,8). The studies in post-mortem brain tissues of sporadic

Alzheimer's patients have shown that UDG expression levels are significantly reduced and also the recognition of uracil lesion and incision activities of UDG decreases (9). In these studies, UDG protein levels and uracil incision activities were found to be low both in inferior parietal lobule (the most affected region) and in the cerebellum (less affected region). The presence of BER lesions due to deficiency in UDG activity in post-mortem brain tissue of patients diagnosed with mild cognitive impairment, which is considered to be a transition period between normal aging and dementia, indicates that defects in BER mechanism can also be observed in the early onset of sporadic AD. This suggests that BER pathogenesis is not limited to neuropathological brain regions, but also it is general character of Alzheimer brain regions (9–13).

Therefore, in this thesis study, *UNG* gene is selected as a genetic risk factor candidate for the sporadic AD. It is also investigated whether the significant decrease in the protein level, expression and activity of UDG in sporadic AD patients is due to genetic variant(s) in *UNG* gene. Thus, *UNG* gene (promoter, exon and intron regions) in sporadic AD patients and healthy controls were sequenced using Ion-PGM NGS technology and the significant *UNG* gene variants were confirmed with Sanger sequencing. Twenty four *UNG* gene variants identified from Ion-PGM sequencing experiments were selected according to their Fisher exact test p values for Sanger sequencing confirmation. Of the 24 variants, 12 *UNG* gene variants were found to be false positive (Table 4.2) and 6 variants were found to be positive (Table 4.3). Thus, the accuracy of NGS variants especially INDELs must be confirmed with Sanger sequencing. The confirmed variants of *UNG* gene are rs1018782, rs1018783, rs1610925, rs2430678, rs2515913 and rs2268406 (Table 4.4). Rs1018782 and rs1018783 are located at the upstream (promoter) of *UNG1* gene; rs1610925 are located at the second intron of *UNG1* and third intron of *UNG2* genes; rs2430678, rs2515913 and rs2268406 are located at the fifth intron of *UNG1* and sixth intron of *UNG2* genes (Figure 4.11). The results of this study demonstrated that there is a significant association between *UNG*/rs1610925 and sporadic AD risk (odds ratio: 2.330; 95% CI=1.069-5.079; p=0.020). In addition, according to the allele and genotype frequencies, rs2268406 can also be considered as a significant difference when the sporadic AD patients are compared with healthy control groups (odds ratio:

2.028; 95% CI=0.928-4.434; p=0.049). Increasing the number of healthy control groups in the study will make this SNPs situation clearer. On the other hand, there is no significant association between rs1018782, rs1018783, rs2430678, and rs2515913 and the sporadic AD patients (p>0.05) (Table 4.5).

Among six *UNG* variants, only rs1018783 is studied before, and rs1018783 is linked to the lung cancer risk among non-Hispanic white smokers (96). This is the first study demonstrating the association between rs1610925 (T/TTA insertion at chr12: 109,537,423^109,537,424) and rs2268406 (T/G SNP at chr12:109,546,886) and sporadic AD risk. Because AD is a complex disease, the epistasis relationship among six *UNG* variants were analyzed, but did not find any significant epistasic effects related with sporadic AD.

The effects on these two *UNG* variants on UDG function are not known yet. Moreover, no other reported genetic variants in the intronic region of *UNG* gene (dsSNP database) where rs1610925 and rs2268406 located were studied for their effect in UDG function. In this study, any genetic variant identified in coding regions of *UNG* gene was not significant. Thus, this study also demonstrated that the decrease in the expression, protein and activity levels of UDG protein in AD is not a consequence of genetic variants in *UNG* gene. That should be due to the other regulatory mechanisms that affect the UDG protein.

The *UNG* gene covering promoter, exon and intron regions was sequenced with NGS technology in the Turkish population for the first time with this study. Independent from sporadic AD outcome, there are common *UNG* variants were identified in all 248 samples (Table 4.8). For example, rs246078 were determined in 97.6% of all the samples (94.78%-G/G homozygous and 2.82% A/G heterozygous SNP). There is no report representing the function of this SNP. Another example is rs2541886 were determined in 70.97% of all the samples (28.63%-T/T homozygous and 42.34% G/T heterozygous SNP). It was reported that rs2541886 significantly correlate with virus diversity (p=0.003) (97). *UNG* is one of the immune responses gene that interacts with human immunodeficiency virus type 1 (HIV-1) infection products and help to degradation of nascent HIV-1 DNA (97). The presence of HIV-1 associated rs2541886 *UNG* variant in the sample population of this study indicates

the presence of HIV-1 in this geographic region. Rs246079 is another example to common variation identified 58.06% of the samples. It is demonstrated that rs246079 is a risk factor for rheumatoid arthritis, lung cancer and prostate cancer (98,99). Yin et al showed that rs246079 is related to decreased risk of esophageal cancer in Chinese population (100).

In conclusion, this study revealed *UNG*/rs1610925 as a new genetic risk factor for sporadic AD. There is a strong association between *UNG*/rs1610925 and sporadic AD risk. In addition, *UNG*/rs2268406 can also be considered as the sporadic AD risk factor. Furthermore, this study also demonstrated that the decrease in the level of UDG protein amount and gene expression, and UDG activity in AD is not due to a *UNG* gene variant. Thus, the results in this thesis besides revealing new genetic risk factor for sporadic AD it will contribute to understand the molecular mechanism of the BER in sporadic AD as well as in neurodegeneration.

REFERENCES

1. Masters CL, Bateman R, Blennow K, Rowe CC, Sperling RA, Cummings JL. Alzheimer's disease. Nature Reviews Disease Primers. 2015;1:1–18.
2. Hou Y, Song H, Croteau DL, Akbari M, Bohr VA. Genome instability in Alzheimer disease. Mechanisms of Ageing and Development. 2017;161:83–94.
3. Krokan HE, Bjøra M. Base Excision Repair. 2013;XXXIII(2):81–7. Available at:
4. 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 12q22-q24.1. Genomics. 1996;36:408–16.
5. Otterlei M, Haug T, Nagelhus TA, Slupphaug G, Lindmo T, Krokan HE. Nuclear and mitochondrial splice forms of human uracil-DNA glycosylase contain a complex nuclear localisation signal and a strong classical mitochondrial localisation signal, respectively. Nucleic Acids Research. 1998;26(20):4611–7.
6. Kruman II, Schwartz E, Kruman Y, Cutler RG, Zhu X, Greig NH, vd. Suppression of uracil-DNA glycosylase induces neuronal apoptosis. Journal of Biological Chemistry. 2004;279(42):43952–60.
7. Ravaglia G, Forti P. Homocysteine and folate as risk factors for dementia and Alzheimer disease. Am J Clin Nutr. 2005;82(May):626–43.
8. Kruman II, Kumaravel TS, Lohani A, Pedersen WA, Cutler RG, Kruman Y, Haughey N, Lee J, Evans M, Mattson MP. 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.
9. Weissman L, Jo D-G, Sorensen MM, de Souza-Pinto NC, Markesbery WR, Mattson MP, Bohr VA. Defective DNA base excision repair in brain from individuals with Alzheimer's disease and amnestic mild cognitive impairment. Nucleic Acids Research. 2007;35(16):5545–55.
10. Lovell MA, Markesbery WR. Oxidative Damage in Mild Cognitive Impairment and Early Alzheimer's Disease. Journal of neuroscience research. 2007;85:3036–40.
11. Yang JL, Weissman L, Bohr VA, Mattson MP. Mitochondrial DNA damage and repair in neurodegenerative disorders. DNA Repair. 2008;7(7):1110–20.
12. Jeppesen DK, Bohr VA., Stevnsner T. DNA Repair Deficiency in Neurodegenaration. DNA Repair. 2011;94(2):166–200.
13. Swain U, Rao KS. Age-dependent decline of DNA base excision repair activity in rat cortical neurons. Mechanisms of Ageing and Development. 2012;133(4):186–94.
14. Bettens K, Sleegers K, Van Broeckhoven C. Genetic insights in Alzheimer's disease. The Lancet Neurology. 2013;12(1):92–104.
15. Handel AE, Disanto G, Ramagopalan S V. Next-generation sequencing in understanding complex neurological disease. Expert Review of Neurotherapeutics. 2013;13(2):215–27.
16. Hunter S, Brayne C. Understanding the roles of mutations in the amyloid precursor protein in Alzheimer disease. Molecular Psychiatry. 2017;23(1):81–93.

17. Martorana A, Esposito Z, Koch G. Beyond the cholinergic hypothesis: Do current drugs work in Alzheimer's disease? CNS Neuroscience and Therapeutics. 2010;16(4):235–45.
18. Mudher A, Lovestone S. Alzheimer's disease - Do tauists and baptists finally shake hands? Trends in Neurosciences. 2002;25(1):22–6.
19. Hardy J, Allsop D. Amyloid deposition as the central event in the aetiology of Alzheimer's disease. Trends in Pharmacological Sciences. 1991;12(C):383–8.
20. Hunter S, Brayne C. Relationships between the amyloid precursor protein and its various proteolytic fragments and neuronal systems. Alzheimer's Research and Therapy. 2012;4(2).
21. Jack CR, Bennett DA, Blennow K, Carrillo MC, Feldman HH, Frisoni GB, Hampel H, Jagust WJ, Johnson KA, Knopman DS, Petersen RC, Scheltens P, Sperling RA, Dubois B. A new classification system for AD , independent of cognition A / T / N : An unbiased descriptive classification scheme for Alzheimer disease biomarkers. Neurology. 2016;0(July):1–10.
22. Chapman KR, Bing-Canar H, Alosco ML, Steinberg EG, Martin B, Chaisson C, Kovwall N, Tripodis Y, Stem RA. Mini Mental State Examination and Logical Memory scores for entry into Alzheimer's disease trials. Alzheimer's Research and Therapy. 2016;8(1):1–11.
23. Siddiqui T, Zia MK, Ali SS, Rehman AA, Ahsan H, Khan FH. Reactive oxygen species and anti-proteinases. Archives of Physiology and Biochemistry. 2016;122(1):1–7.
24. Tubbs A, Nussenzweig A. Endogenous DNA Damage as a Source of Genomic Instability in Cancer. Cell. 2017;168(4):644–56.
25. Kasai H. What causes human cancer? Approaches from the chemistry of DNA damage. Genes and Environment. 2016;38(1):1–13.
26. Georgakilas AG, O'Neill P, Stewart RD. Induction and Repair of Clustered DNA Lesions: What Do We Know So Far? Radiation Research. 2013;180(1):100–9.
27. Kunkel TA, Erie DA. Eukaryotic Mismatch Repair in Relation to DNA Replication. Annual Review of Genetics. 2015;49(1):291–313.
28. Liu D, Keijzers G, Rasmussen LJ. DNA mismatch repair and its many roles in eukaryotic cells. Mutation Research. 2017;773:174–87.
29. Selzer RR, Nyaga S, Tuo J, May A, Muftuoglu M, Christiansen M, Citterio E, Brosh RM, Bohr VA. Differential requirement for the ATPase domain of the Cockayne syndrome group B gene in the processing of UV-induced DNA damage and 8-oxoguanine lesions in human cells. Nucleic acids research. 2002;30(3):782–93.
30. Muftuoglu M, Sharma S, Thorslund T, Stevnsner T, Soerensen MM, Brosh RM, Bohr VA. Cockayne syndrome group B protein has novel strand annealing and exchange activities. Nucleic Acids Research. 2006;34(1):295–304.
31. Yu S, Evans KE, van Eijk P, Bennett M, Webster RM, Leadbitter M, Teng Y, Waters R, jackson SP, Reed SH. Global Genome Nucleotide Excision Repair is Organised into Domains Promoting Efficient DNA Repair in Chromatin. Genome Research. 2016;26:1376–87.
32. de Laat W, Jaspers N, Hoeijmakers J. Molecular mechanism of nucleotide excision repair. Genes & Development. 1999;13(7):768–85.

33. Carter RJ, Parsons JL. Base Excision Repair , a Pathway Regulated by Posttranslational. 2016;36(10):1426–37.
34. Prakash A, Doublie S. Base Excision Repair in the Mitochondria. 2015;116(8):1490–9.
35. Thapar U, Demple B. How are base excision DNA repair pathways deployed in vivo? F1000Research. 2017;6.
36. Shevelev I V., Hübscher U. The 3′-5′ exonucleases. Nature Reviews Molecular Cell Biology. 2002;3(5):364–75.
37. De Ory A, Nagler K, Carrasco B, Raguse M, Zafra O, Moeller R, vd. Identification of a conserved 5′-dRP lyase activity in bacterial DNA repair ligase D and its potential role in base excision repair. Nucleic Acids Research. 2016;44(4):1833–44.
38. Muftuoglu M, Wong HK, Imam SZ, Wilson DM, Bohr VA, Opresko PL. Telomere repeat binding factor 2 interacts with base excision repair proteins and stimulates DNA synthesis by DNA polymerase β. Cancer Research. 2006;66(1):113–24.
39. Barzilai A, Schumacher B, Shiloh Y. Genome instability: Linking ageing and brain degeneration. Mechanisms of Ageing and Development. 2017;161:4–18.
40. Swain U, Subba Rao K. Study of DNA damage via the comet assay and base excision repair activities in rat brain neurons and astrocytes during aging. Mechanisms of Ageing and Development. 2011;132(8–9):374–81.
41. Canugovi C, Misiak M, Ferarelli LK, Croteau DL, Bohr VA. The role of DNA repair in brain related disease pathology. DNA Repair. 2013;12(8):578–87.
42. Friedberg EC. A history of the DNA repair and mutagenesis field. The discovery of base excision repair. DNA Repair. 2016;37:A35–9.
43. Madabhushi R, Pan L, Tsai LH. DNA damage and its links to neurodegeneration. Neuron. 2014;83(2):266–82.
44. Lu T, Pan Y, Kao S, Li C, Kohane I, Chan J, Yankner BA. Gene regulation and DNA damage in the ageing human brain. Nature. 2004;429(6994):883–91.
45. Bahar R, Hartmann CH, Rodriguez KA, Denny AD, Bussuttil RA, Dolle MET, vdCalder RB, Chisholm GB, Pollock BH, Klein CA, Vijg J. Increased cell-to-cell variation in gene expression in ageing mouse heart. Nature. 2006;441:1011–23.
46. Vermeij WP, Hoeijmakers JHJ, Pothof J. Aging: not all DNA damage is equal. Current Opinion in Genetics & Development. 2014;26:124–30.
47. Kadowaki H, Nishitoh H, Urano F, Sadamitsu C, Matsuzawa A, Takeda K, Masutani H, Yodof J, Urano Y, Nagano T, Ichijo H. Amyloid β induces neuronal cell death through ROS-mediated ASK1 activation. Cell Death and Differentiation. 2005;12(1):19–24.
48. Markesbery WR, Lovell MA. DNA Oxidation in Alzheimer ' s Disease. Antioxidants and Redox Signaling. 2006;8(11–12):2039–45.
49. Akbari M, Morevati M, Croteau D, Bohr VA. The role of DNA base excision repair in brain homeostasis and disease. DNA Repair. 2015;32:172–9.
50. Birch-Machin MA, Bowman A. Oxidative stress and ageing. British Journal of Dermatology. 2016;175:26–9.
51. Mosch B, Morawski M, Mittag A, Lenz D, Tarnok A, Arendt T. Aneuploidy and DNA Replication in the Normal Human Brain and Alzheimer's Disease. Journal of

Neuroscience. 2007;27(26):6859–67.

52. Mullaart E, Boerrigter M, Ravid R, Waab D, Vijg J. Increased Levels of DNA Breaks i Cerebral Cortex of Alzheimer's Disease Patients. C. 11, Neurobiology of Aging. 1990. s. 169–73.
53. Suberbielle E, Sanchez PE, Kravitz A V., Wang X, Ho K, Eilertson K, Devidze N, Kreitzer AC, Mucke L. Physiologic brain activity causes DNA double-strand breaks in neurons, with exacerbation by amyloid-β. Nature Neuroscience. 2013;16(5):613–21.
54. Bradley-Whitman MA, Timmons MD, Beckett TL, Murphy MP, Lynn BC, Lovell MA. Nucleic acid oxidation: An early feature of Alzheimer's disease. Journal of Neurochemistry. 2014;128(2):294–304.
55. Lovell MA, Markesbery WR. Oxidative DNA damage in mild cognitive impairment and late-stage Alzheimer's disease. Nucleic Acids Research. 2007;35(22):7497–504.
56. Lovell MA, Somana S, Bradley MA. oxydatively modified nucleic acids in preclinical Alzheimer's disease (PCAD) brain. mech aging dev. 2011;132(8–9):443–8.
57. Coppedè F, Tannorella P, Stoccoro A, Chico L, Siciliano G, Bonuccelli U, Migliore L. Methylation analysis of DNA repair genes in Alzheimer's disease. Mechanisms of Ageing and Development. 2017;161:105–11.
58. Zawia NH, Lahiri DK, Cardozo-Pelaez F. Epigenetics, oxidative stress, and Alzheimer disease. Free Radical Biology and Medicine. 2009;46(9):1241–9.
59. Sykora P, Misiak M, Wang Y, Ghosh S, Leandro GS, Liu D, vd. DNA polymerase β deficiency leads to neurodegeneration and exacerbates Alzheimer disease phenotypes. Nucleic Acids Research. 2015;43(2):943–59.
60. Dizdaroglu M, Jaruga P, Birincioglu M, Rodriguez H. Free radical-induced damage to DNA: Mechanisms and measurement. Free Radical Biology and Medicine. 2002;32(9):804–12.
61. Lovell MA, Gabbita SP, Markesbery WR. Increased DNA oxidation and decreased levels of repair products in Alzheimer's disease ventricular CSF. Journal of Neurochemistry. 1999;72(2):771–6.
62. Sobol RW, Horton JK, Kühn R, Gu H, Singhal RK, Prasad R, vd. Requirement of mammalian DNA polymerase-β in base-excision repair. Nature. 1996;379(11):183–6.
63. Kwiatkowski D, Czarny P, Galecki P, Bachurska A, Talarowska M, Orzechowska A, Bobinska K, Bielecka-Kovalska A, Pietras T, Szemraj J, Maes M, Sliwinski T. Variants of base excision repair genes MUTYH, PARP1 and XRCC1 in Alzheimer's disease risk. Neuropsychobiology. 2015;71(3):176–86.
64. Czarny P, Kwiatkowski D, Toma M, Kubiak J, Sliwinska A, Talarowska M, Szemraj J, Maes M, Galecki P, Sliwinski T. Impact of Single Nucleotide Polymorphisms of Base Excision Repair Genes on DNA Damage and Efficiency of DNA Repair in Recurrent Depression Disorder. Molecular Neurobiology. 2017;54(6):4150–9.
65. Kwiatkowski D, Czarny P, Toma M, Jurkowska N, Sliwinska A, Drzewoski J, Bachurska A, Szemraj J, Maes M, Berk M, Su KP, Galecki P, Sliwinski T. Associations between DNA damage, DNA base excision repair gene variability and Alzheimer's disease risk. Dementia and Geriatric Cognitive Disorders.

2016;41(3–4):152–71.

66. D'Errico M, Parlanti E, Pascucci B, Fortini P, Baccarini S, Simonelli V, Dogliotti E. Single nucleotide polymorphisms in DNA glycosylases: From function to disease. Free Radical Biology and Medicine. 2017;107(October 2016):278–91.
67. Norjmaa B, Tulgaa K, Saitoh T. Base Excision Repair Pathway and Polymorphisms of XRCC1 Gene. Journal of Molecular Pathological Epidemiology. 2016;1(4):1–5.
68. Pearl LH. Structure and function in the uracil-DNA glycosylase superfamily. Mutation Research - DNA Repair. 2000;460(3–4):165–81.
69. Olsen LC, Aasland R, Wittwer CU, Krokan HE, Helland DE. Molecular cloning of human uracil DNA glycosylase, a highly conserved DNA repair enzyme. The EMBO Journal. 1989;8(10):3121–5.
70. Varshney U, Hutcheon T, van de Sande JH. Sequence analysis, expression, and conservation of Escherichia coli uracil DNA glycosylase and its gene UNG. J Biol Chem. 1988;263(16):7776–84.
71. Kavli B, Andersen S, Otterlei M, Liabakk NB, Imai K, Fischer A, Durandy A, Krokan HE, Slupphaug G. B cells from hyper-IgM patients carrying *UNG* mutations lack ability to remove uracil from ssDNA and have elevated genomic uracil. The Journal of Experimental Medicine. 2005;201(12):2011–21.
72. Nilsen H, Otterlei M, Haug T, Solum K, Nagelhus TA, Skorpen F, Krokan HE. Nuclear and mitochondrial uracil-DNA glycosylases are generated by alternative splicing and transcription from different positions in the UNG gene. Nucleic Acids Research. 1997;25(4):750–5.
73. Lund H, Eftedal I, Haug T, Krokan HE. Pseudogenes for the human uracil-DNA glycosylase on chromosomes 14 and 16. Biochemical and Biophysical Research Communications. 1996;224(1):265–70.
74. Xia B, Liu Y, Guevara J, Li J, Jilich C, Yang Y, Wang L, Dominy BN, Cao W. Correlated Mutation in the Evolution of Catalysis in Uracil DNA Glycosylase Superfamily. Scientific Reports. 2017;7:1–12.
75. Schormann N, Ricciardi R, Chattopadhyay D. Uracil-DNA glycosylases - Structural and functional perspectives on an essential family of DNA repair enzymes. Protein Science. 2014;23(12):1667–85.
76. Dingler FA, Kemmerich K, Neuberger MS, Rada C. Uracil excision by endogenous SMUG1 glycosylase promotes efficient Ig class switching and impacts on A:T substitutions during somatic mutation. European Journal of Immunology. 2014;44(7):1925–35.
77. Bharati S, Krokan HE, Kristiansen L, Otterlei M, Slupphaug G. Human mitochondrial uracil-DNA glycosylase preform (UNG1) is processed to two forms one of which is resistant to inhibition by AP sites. Nucleic Acids Research. 1998;26(21):4953–9.
78. Rodriguez G, Esadze A, Weiser BP, Schonhoft JD, Cole PA, Stivers JT. Disordered N-Terminal Domain of Human Uracil DNA Glycosylase (hUNG2) Enhances DNA Translocation. ACS Chemical Biology. 2017;12(9):2260–3.
79. Parikh SS, Mol CD, Slupphaug G, Bharati S, Krokan HE, Tainer JA. Base excision repair initiation revealed by crystal structures and binding kinetics of human uracil-DNA glycosylase with DNA. EMBO Journal. 1998;17(17):5214–26.
80. Parikh SS, Walcher G, Jones GD, Slupphaug G, Krokan HE, Blackburn GM,

Tainer JA. Uracil-DNA glycosylase-DNA substrate and product structures: conformational strain promotes catalytic efficiency by coupled stereoelectronic effects. Proceedings of the National Academy of Sciences of the United States of America. 2000;97(10):5083–8.

81. Nilsen H, Rosewell I, Robins P, Skjelbred CF, Andersen S, Slupphaug G, Daly G, Krokan HE, Lindahl T, Barnes DE. Uracil-DNA Glycosylase (UNG)-Deficient Mice Reveal a Primary Role of the Enzyme during DNA Replication. Molecular Cell. 2000;5:1059–65.
82. Imai K, Slupphaug G, Lee WI, Revy P, Nonoyama S, Catalan N, Yel L, Forveille M, Kavli B, Krokan HE, Ochs HD, Fischer A, Durandy A. Human uracil-DNA glycosylase deficiency associated with profoundly impaired immunoglobulin class-switch recombination. Nature Immunology. 2003;4(10):1023–8.
83. Rada C, Williams GT, Nilsen H, Barnes DE, Lindahl T, Neuberger MS. Immunoglobulin isotype switching is inhibited and somatic hypermutation perturbed in UNG-deficient mice. Current Biology. 2002;12(20):1748–55.
84. Miller DT, Adam MP, Aradhya S, Biesecker LG, Brothman AR, Carter NP, Church DM, Crolla JA, Eicher EE, Epstein CJ, Faucett EB, Feuk L, Friedman JM, Hamosh A, Jackson L, Kaminsky EB, Kok L. Kuhn RM, Lee C, Ostell JM, Rosenberg C, Scherer SW, Spinner NB, Stravropoulos DJ, Tepperberg JH, Thorland EC, Vermeesch JR, Waggoner DJ, Watson MS, Martin CL, Ledbetter DH. Consensus Statement : Chromosomal Microarray Is a First-Tier Clinical Diagnostic Test for Individuals with Developmental Disabilities or Congenital Anomalies. The American Journal of Human Genetics. 2010;86(5):749–64.
85. Al-hertani W, Mcgowan-Jordan J, Allanson JE. Novel Clinical Findings in a Case of Postnatally Diagnosed Trisomy 12 Mosaicism. american journal of medical genetics. 2012;158:1452–4.
86. Roberts W, Zurada A, Nska AZ, Gielecki J, Loukas M. Anatomy of Trisomy 12. clinical anatomy. 2016;29:633–7.
87. Kaminsky EB, Kaul V, Paschall J, Church DM, Bunke B, Kunig D, Moreno-D-Luca D, Moreno-De-Luca A, Mulle JG, Warren ST, Richard G, Compton JG, Fuller AE, Gliem TJ, Huang S, Collinson MN, Beal SJ, Ackley T, Pickering DL, Golden DM, Aston E, Whitby H, Shetty S, Rossi MR, Rudd MK, South ST, Brothman AR, Sanger WG, Iyer RK, Crolla JA, Thorland EC, Aradhya S, Ledbetter DH, Martin CL. An evidence-based approach to establish the functional and clinical significance of CNVs in intellectual and developmental disabilities. Genetics in Medicine. 2011;13(9):777–84.
88. Xue JH, Xu GF, Gu TP, Chen GD, Han B Bin, Xu ZM, bjoras M, Krokan HE, Xu GL, Du YR. Uracil-DNA glycosylase UNG promotes tet-mediated DNA demethylation. Journal of Biological Chemistry. 2016;291(2):731–8.
89. Ma'ayan A. Complex systems biology. Journal of the Royal Society Interface. 2017;14(134).
90. Mardis ER. Next-Generation DNA Sequencing Methods. Annual Review of Genomics and Human Genetics. 2008;9(1):387–402.
91. Metzker ML. Sequencing technologies the next generation. Nature Reviews Genetics. 2010;11(1):31–46.
92. İlhan T. Sporadik Alzheimer Hastaliğinda DNA Hedef Gen Varyantlarinin Taranmasi. Acibadem University, Sağlık Bilimleri Enstitüsü, Yüksek Lisans Tezi, İstanbul, 2016.

93. Herron JC, Hodin J. Evolutionary Analysis. Fifth edition. Pearson. 2007.

94. Szumilas M. Explaining odds ratios. Journal of the Canadian Academy of Child and Adolescent Psychiatry. 2010;19(3):227–9.

95. Pearson T a, Manolio TA. How to interpret a genome-wide association study. JAMA : the journal of the American Medical Association. 2008;299(11):1335–44.

96. Doherty JA, Sakoda LC, Loomis MM, Barnett MJ, Julianto L, Thornquist MD, Neuhouser ML, Weiss NS, Goodman GE, Chen C. DNA Repair genotype and lung cancer risk in the betacarotene and retinol efficacy trial. International Journal of Molecular Epidemiology and Genetics. 2013;4(1):11–34.

97. Fumagalli M, Pozzoli U, Cagliani R, Comi GP, Bresolin N, Clerici M, Sironi M. Genome-wide identification of susceptibility alleles for viral infections through a population genetics approach. PLoS Genetics. 2010;6(2).

98. De Mars G, Windelinckx A, Huygens W, Peeters MW, Beunen GP, Aerssens J, Vlietinck R, Thomis MAI. Genome-wide linkage scan for maximum and length-dependent knee muscle strength in young men: Significant evidence for linkage at chromosome 14q24.3. Journal of Medical Genetics. 2008;45(5):275–83.

99. Barry KH, Koutros S, Andreotti G, Sandler DP, Burdette LA, Yeager M, Freeman LEB, Lubin JH, Ma X, Zheng T, Alavanja MCR, Berndt SI. Genetic variation in nucleotide excision repair pathway genes, pesticide exposure and prostate cancer risk. Carcinogenesis. 2011;33(2):331–7.

100. Yin J, Sang Y, Zheng L, Wang L, Yuan L, Liu C, Wang X, Shi Y, Shao A, Ding G, Chen S, Tang W, Gu H. Uracil-DNA glycosylase (UNG) rs246079 G/A polymorphism is associated with decreased risk of esophageal cancer in a Chinese population. Medical Oncology. 2014;31(11).

CURRICULUM VITAE

Personal Information

Name	Aslı	**Surname**	Semerci
Birth of place	Ankara	**Date of birth**	01.10.1993
Nationality	Turkish	**Telephone number**	905377223848
E-mail	semerciasli@gmail.com		

Educational Level		
	Institution Name	**Graduate Year**
Master	Mehmet Ali Aydınlar Acibadem University	2018
Undergraduate	Istanbul University	2015
High-School	Nermin Mehmet Çekiç Anadolu Lisesi	2011

Work Experience

	Position	**Enterprise**	**Duration**
1.	Founder-CEO	GeneCare Biotechnology	2016-
2.	Laboratory Technician	Mavi Medical Laboratory	2013-2015

Foreign Languages	**Reading***	**Speaking***	**Writing***
English	Very good	Good	Good
German	Weak	Weak	Weak

*Evaluated with very good, good, medium and weak

Foreign Languages Exams•

KPDS	**UDS**	**IELTS**	**TOEFL IBT**	**TOEFL PBT**	**TOEFL CBT**	**FCE**	**CAE**	**CPE**

•All of the successful exams should be enrolled.
KPDS: Kamu Personeli Yabancı Dil Sınavı; ÜDS: Üniversitelerarası Kurul 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

	Quantitative	**Equally Weighted**	**Verbal**
ALES Exam	85,8	87,3	74,6
(Other) Exam			

Projects				
	Granted Program	**Active Years**	**Position**	**Project Name**
1.	TUBITAK 1512	2016-2018	Coordinator	Improving a methodology for early diagnosis of Alzheimer's disease.
2.	TUBITAK 1001	2015-2018	Scholarship student	Investigation of the variants of base excision repair genes for sporadic Alzheimer's disease.
3.	TUBITAK 2209/A	2014-2015	Coordinator	Investigation of Arabidopsis thaliana GIGANTUS1 (GTS1) expression level on the different stress conditions.
4.	KOSGEB	2013-2014	Researcher	Microarray chip design for detection of genetically modified organisms

Presentations & Awards			
	Organization	**Presentation Type**	**Title of the Presentation**
1.	Proje Park'15 Biyo-Med	Poster	Early diagnosis of Alzheimer's disease with tracking of membrane metalloproteinases.
2.	Proje Park'15 Biyo-Med	Poster	Targeted therapy for Alzheimer's disease.
3.	9. Aykut Kence Evolution Conference	Oral	The Epidemiology of HIV
4.	1st HBP School	Poster	β-Amyloid plaque production in HIV infected individuals.
5.	TEB-Let’s Up ‘16	-	Grant award