A CANDIDATE GENE STUDY OF LATE-ONSET ALZHEIMER'S DISEASE

by

Lauren Christina Burns

BS, California Polytechnic State University, 2007

Submitted to the Graduate Faculty of

the Department of Human Genetics of the

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2010

UNIVERSITY OF PITTSBURGH

Graduate School of Public Health

This thesis was presented

by

Lauren Christina Burns

It was defended on

April 13, 2010

and approved by

Thesis Advisor:
M. Ilyas Kamboh, Ph.D.
Professor and Chair
Department of Human Genetics
Graduate School of Public Health
University of Pittsburgh

Thesis Committee Member:
David Finegold, MD, Ph.D.
Professor
Department of Pediatrics and Department of Human Genetics
School of Medicine and Graduate School of Public Health
University of Pittsburgh

Thesis Committee Member:
M. Michael Barmada, Ph.D.
Associate Professor
Department of Human Genetics
Graduate School of Public Health
University of Pittsburgh

A CANDIDATE GENE STUDY OF LATE-ONSET ALZHEIMER'S DISEASE

Lauren Christina Burns, M.S.

University of Pittsburgh, 2010

Late-onset Alzheimer's disease (LOAD) is a multifactorial disease with the potential involvement of multiple genes. Complex diseases such as LOAD have a large affect on public health due to the significant financial burden of health care for these individuals. Genetics plays a significant role in the etiology of the disease, therefore it is of public health importance that the genetics of LOAD be investigated. There is a known association between *APOE* gene variants and LOAD. No additional genes have been consistently demonstrated to be associated with risk of LOAD. Multiple recent genome-wide association studies (GWAS) have found variants showing significant association with LOAD. Twelve SNPs were chosen from four GWAS for a replication study to determine if the associations seen between SNPs and AD risk in the respective studies were present in our population. Ten additional positional candidate SNPs were chosen on chromosome 10 because of observed linkage peaks for AD and predicted imprinted genes in this region. We genotyped these 22 SNPs as well the *E2*/*E3*/*E4 APOE* polymorphism in up to 993 Caucasian Americans with LOAD and up to 976 age-matched healthy Caucasian Americans. Our data showed no statistically significant associations between the 22 SNPs examined and the risk of AD. Stratification by *APOE*4* carrier status also failed to reveal statistically significant associations. Additional analyses were performed to examine the potential associations between the 22 SNPs and age-at-onset, disease duration, and baseline MMSE score. The analysis revealed significant associations between two SNPs, rs3746319 (p=0.002) and rs16934131 (p=0.045), and age-at-onset. One SNP, rs16934131 (p=0.0002), was

found to be significantly associated with disease duration. Three SNPs, rs16934131 (p=0.002), rs12781609 (p=0.012), and rs5984894 (p=0.009), were found to be associated with baseline MMSE score in controls. Of note, rs16934131, demonstrated statistically significant association with age-at-onset, disease duration, and MMSE score. Further study may be necessary to definitively rule out associations between these variants and LOAD.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

First and foremost I would like to thank Dr. M. Ilyas Kamboh for giving me the opportunity to work in his lab. The experience and knowledge that I have gained during my time working for this lab are invaluable. To my other committee members, Dr. M. Michael Barmada and Dr. David Finegold, thank you for your support and guidance through the process of my masters thesis. Additionally, I would like to thank Dr. Eleanor Finegold, Dr. Candace Kammerer, and Dr. Daniel Weeks, for their guidance in the statistical analysis of my thesis.

I would also like to thank my co-workers in the lab who have provided their moral support and knowledge when I needed it the most. It has been a pleasure to work with you. I would also like to thank Zahra Rahman for her moral and emotional support during my time in the lab. I would also like to thank Dr. Oussama Khalifa for his help with lab work, which was always done with a smile. I would especially like to thank Ryan Minster for his patience, guidance, and support throughout the entire thesis process.

I would also like to express my gratitude to the directors of the Genetic Counseling program, Elizabeth Gettig and Dr. Robin Grubs. You have both taught me so much and I thank you for giving me the opportunity to grow professionally and personally.

Finally, I would like to thank my family and friends, who have loved and supported me during my time at the University of Pittsburgh. Without you I would not have been able to accomplish all that I have.

1.0 BACKGROUND

1.1 INTRODUCTION

Alzheimer's disease (AD) is the most common neurodegenerative disease and accounts for approximately two-thirds of individuals with dementia. AD can be classified by the age-at-onset (AAO) and the presence or lack of family history. AD can be divided into early-onset (before 60 years of age) and late-onset (after 60 years of age). Sporadic AD is classified as a lack of significant family history of AD and familial AD is classified by a significant family history. Sporadic AD makes up approximately 90% of cases, while familial AD accounts for the remaining 10%. Three genes have been implicated in the etiology of early-onset AD: amyloid precursor protein (*APP*), presenilin 1 (*PSEN1*), and presenilin 2 (*PSEN2*). These genes follow an autosomal dominant pattern of inheritance and represent approximately 50% of early-onset AD and less than 1% of all AD cases. Late-onset AD (LOAD) is a thought to be a multifactorial disease, and is therefore potentially affected by multiple genes and environmental factors. To date, the only known significant genetic risk factor for LOAD is apolipoprotein E (*APOE)* (Kamboh 2004).

1.2 EPIDEMIOLOGY OF ALZHEIMER'S DISEASE

In 2009, it was estimated that 5.3 million people in the United States had AD, of which 5.1 million were age 65 or older (Alzheimer's Association 2009). Worldwide, about 35 million people are affected with AD (World Alzheimer Report 2009). The risk of AD increases with age, therefore the number of persons with AD is expected to grow as more individuals continue to live longer. Due to recent advances in medicine enabling more individuals to live into their 80s and 90s, the burden of AD has dramatically increased over the past decades and continues to rise. In 2000, there were an estimated 411,000 new cases of AD in the United States and this number is expected to increase to 454,000 new cases in 2010. By 2050, the number of new cases of AD is projected to be 959,000 (Alzheimer's Association 2009).

In 2009, the total annual cost of Alzheimer's and other dementias in the United States was estimated to be 148 billion dollars. Individuals with AD and other dementias are high users of healthcare and long-term care services. This includes hospital stays due to disease-related injuries and complications, as well as adult day center services, assisted living, and nursing home care (Alzheimer's Association 2009). The average individual with AD requires 2.75 years of nursing home care, over 10 times the national median length of stay for all causes (Welch et al. 1992). A prolonged period of institutionalization and the high expensive of long term care combine to make the costs of AD extensive. The great proportion of the financial burden in the elderly population and greater is placed on Medicare and/or Medicaid. In 2004, the total per-person payments for health and long-term care were three times higher for Medicare beneficiaries over 65 years with AD than for other Medicare beneficiaries in the same age group (Alzheimer's Association 2009). The costliness of all medical services necessary to care for the ever-growing AD population add up to a tremendous financial burden.

1.3 PATHOGENESIS AND DIAGNOSIS

There are two pathological features of AD: extracellular senile plaques and the presence of intracellular neurofibrillary tangles. These extracellular senile plaques are mainly made up of the 42 amino acid amyloid β peptide (Aβ-42) that is derived from the amyloid precursor protein (APP). APP is a transmembrane protein that undergoes proteolytic cleavage by either α- or β-secretase and subsequent cleavage by γ-secretase. When APP is cut by α-secretase and then by γ-secretase, a harmless peptide is generated. Conversely, when APP is cut first by β-secretase and then by γ-secretase, it generates peptides that vary between 39-43 amino acids in length. Aβ-40 is the most common and Aβ-42 accounts for approximately 10% of Aβ. Aβ-40 is considered harmless, while Aβ-42 is neurotoxic and involved in the formation of senile plaques in AD brains (see Kamboh 2004 and references therein).

Microtubules are one of the main structural components of cells. Microtubules are made up of a protein called tubulin that polymerizes to form straight, thin structures that provide the cell with support, shape, and transportation for substances. The microtubule network is held together by tau, which helps provide stability to the cell. In the brains of AD, hyperphosphorylation causes tau to self-aggregate, leading the microtubule assembly to collapse. The aggregated tau molecules form a twisted paired-helical filament-like structure, thus named tau neurofibrillary tangles (see Kamboh 2004 and references therein).

A clinical diagnosis of AD can be made following a neurological exam that includes various cognitive parameters. These clinical features include loss of the ability to smell, progressive decline in memory and intellectual ability, deteriorating language and speech skills, loss of orientation and behavioral abilities (see Kamboh 2004 and references therein). Clinical

diagnoses can be classified as possible or probable. Probably AD is when the individual has no other illness that may be contributing to symptoms whereas possible AD is when the individual meets criteria for another illness that may contribute to cognitive delay (Vilalta-Franch et al. 2010). A definitive diagnosis of AD can only be made based on post-mortem studies, specifically the presence of pathological hallmarks in the brain.

One method for determining an individual's degree of cognitive impairment and/or dementia is the mini-mental state examination (MMSE). The MMSE measures orientation to time and place, registration, attention and calculation, recall, language, repetition, and complex commands using various types of tests (Folstein et al. 1975).

1.4 ENVIRONMENTAL RISK FACTORS FOR LOAD

LOAD is considered a multifactorial condition in which several environmental and genetic risk factors contribute to the etiology of the disease. Consideration of environmental factors is a necessary part of studying potential genetic factors that may contribute to risk for LOAD. There are several hypothesized environmental risk factors for LOAD, including education level, gender, tobacco and alcohol use, head trauma, cerebrovascular disease, and cholesterol levels. Although there are many possible risk factors for LOAD, most of the data is questionable. The only proven risk factor for LOAD is age.

1.4.1 Education

Several studies have investigated the effect of education level on AD risk. It has been proposed that a low level of education is a risk factor for LOAD. A study by Sando et al. (2008) found that education had a consistently protective influence on the risk of developing AD. They found that this effect was dose-dependent and was still evident after correcting for age, sex, and *APOE*4* status. Other studies found no dose-dependent effect, but rather a threshold effect, in which individuals who had reached a certain education level benefited from the protective effect (Letenneur et al. 1999). The dose-dependent finding by Sando et al. (2008) is important because it suggests that education is causatively protective against AD, rather than a representation of confounding factors.

The mechanism behind the protective effect of education is debated. Pathological studies on brains of individuals with LOAD revealed no correlations between education level and amyloid deposition (Bennett et al. 2005). Therefore, it is hypothesized that higher level of education allows individuals to withstand the effect of amyloid deposition on cognitive function. Therefore, pathologically there is no difference between individuals with varying education levels, but the effect on cognition is changed. “Cognitive reserve” or “brain reserve” theory proposes that education provides a neurological “reserve” through changes in neurons or processing efficiency, which allows the clinical symptoms of neurodegeneration in AD to occur later. Therefore, people with higher levels of education show symptoms later than those with less education (Paradise et al. 2009). Individuals with higher education levels may be more prepared to cope with these pathological changes because their brain can use the pre-exisiting cognitive processing that was acquired throughout their lifetime, particularly during challenging tasks during education (Sando et al. 2008). Cognitive reserve theory also suggests that individuals with

higher levels of education will have shorter disease duration, as they are hypothetically farther in the pathological disease progression at time of their diagnosis. However, Paradise et al. (2009) found that there was no change in time of death after diagnosis based on education. Others theorize that the association between education level and AD risk is not causative, but rather represents a confounding factor. This confounding factor could be a genetic factor representing innate intelligence, which one would infer would lead to a propensity for higher levels of education. This same genetic factor could be responsible for protecting the individual from LOAD. Yet others argue that years of study do not accurately reflect innate intelligence because other factors, such as financial means, may prevent individuals from continuing with schooling. However, the correlation is still seen between higher education levels and lower LOAD risk, suggesting that education, not innate intelligence, is the true protective factor (Sando et al. 2008).

It is important to note that other factors that influence education level may be confounding the data. Individuals who are less educated may have jobs that involve manual labor. These types of occupations may make them more likely to be exposed to toxic substances and head trauma (Letenneur et al. 1999). Additionally, because AD cannot be definitely diagnosed until autopsy, clinical diagnosis is based on the observation and reporting of appearance of symptoms. Lower or higher education levels may affect the time which symptoms are first reported. For example, it may be more difficult to detect symptoms of cognitive impairment in an individual with a lower education level. Alternately, symptoms may be more noticeable in individuals with higher education levels because they may have greater occupational demands (Mejia et al. 2003). All studies that investigate the effect of education on AD risk are at risk for participation bias.

1.4.2 Gender

As there is a difference in LOAD prevalence between men and women, gender is thought to be a risk factor for developing LOAD. One study found incidence of AD to be higher in women than men in all age groups. The risk for dementia was increased by 3.1 in women as compared to men (Letenneur et al. 1999). However, in Letenneur et al.'s study, AD incidence in women was lower than in men before the age of 80 and higher in women than men after age 80, suggesting a difference in AAO. Letenneur et al. (1999) offer several possible explanations for the observed differences in AD risk between men and women. Hormonal differences may explain the delay in occurrence in women. Some studies have found a negative association between estrogen replacement therapy and risk for AD (Kawas et al. 1997). Genetic differences between men and women may also play a part in this observed difference in AD risk between men and women; possible genetic risk factors that may be present on the X chromosome or protective factors on the Y chromosome. While biological factors may account for these differences, it is important to consider confounders such as social differences between men and women. The most prominent cofounder in this category is education level. On average, men have higher levels of education compared to women, particularly in older generations. However, after adjusting for education level, Letenneur et al. (2000) found that the association between AD risk and gender remained statistically significant. Another difference between men and women that could confound the results is alcohol consumption. Moderate wine consumption has been linked to lower AD risk (Letenneur et al. 2004). Therefore it is important to consider that men consume more wine on average than women when assessing differences in AD risk. After adjusting for wine consumption, Letenneur et al. (2004) found that the association between AD and sex remained unchanged. As with any study of this potential association, Letenneur et al. (2004) reported that

participation bias may have affected their results, as most refusals to participate were younger, more often women, and were on average less educated.

1.4.3 Tobacco and Alcohol Usage

Tobacco use has been proposed to have a protective effect against the development of AD. AD is associated with cholinergic system deficits, which is characterized by reduced levels of acetylcholine and nicotinic receptors. Nicotine increases the release of acetylcholine and the number of nicotine receptors, and improves information processing and attention. For this reason, the protective effect of tobacco against AD is biologically plausible. Furthermore, case-control studies have shown that long-term smoking habits are associated with a lower risk of AD in a dose-dependent fashion. Conversely, prospective studies have found an increased risk of AD associated with tobacco use. However, after adjusting for education level and occupation, the association between tobacco use and AD risk was not significant (Letenneur et al. 2004).

Alcohol consumption has also been proposed as having a protective effect against the development of AD. Several studies found that light to moderate alcohol consumption was associated with lower risk of developing AD. The PAQUID study was a prospective study that studied wine consumption specifically, as it was the only type of alcohol consumed daily by most of the subjects (Orgogozo et al. 1997). Subjects were labeled as non-drinkers, light drinkers (1-2 glasses of wine per day), moderate drinkers (3-4 glasses per day), and heavy drinkers (more than four glasses per day). The risk of AD was decreased among light to moderate drinkers when compared to non-drinkers. The odds ratios for light, moderate, and heavy drinkers were 0.55, 0.28, and 0.48, respectively. These results remained unchanged when adjusted for many potential confounders including age, sex, education level, smoking status, and MMSE. These results were

also looked at when stratified by *APOE*4* carriers and non-carriers and showed that a light to moderate consumption seems to reduce the risk of AD among the subjects who have a higher genetic risk of developing AD. The underlying biological explanation for the association between alcohol consumption and AD is still unknown. One theory proposes that alcohol acts through reduction of cardiovascular risk factors, either by inhibiting platelet aggregation or through alteration of the serum lipid profile. Another theory states that alcohol might have a direct effect on cognition through release of acetylcholine in the hippocampus, which would hypothetically facilitate learning and memory (Letenneur et al. 2004).

1.4.4 Head Trauma

The effect of head injury has been related to dementia and AD risk in several studies. A study by Plassman et al. (2000) examined the AD and dementia risk in individuals with documented traumatic brain injury (TBI). TBI was classified as severe (loss of consciousness or posttraumatic amnesia longer than 24 hours), moderate (loss of consciousness of posttraumatic amnesia between 30 minutes and 24 hours), or mild (brief loss of consciousness or posttraumatic amnesia less than 30 minutes). Hazard ratios (HR) were calculated for the severity of head injury and AD. Severe TBI was found to have a HR of 4.51 and moderate TBI had a HR of 2.32. Mild TBI was not found to increase AD risk (HR 0.76). The overall hazard ratio for TBI and AD was 2.0 (Plassman et al. 2000). In 2008, the Institute of Medicine stated that there is sufficient evidence of association between moderate or severe TBI and dementia of the Alzheimer type and that there is limited evidence of an association between mild TBI, with or without loss of consciousness, and dementia (Kumar and Kinsella 2010).

Pathological features similar to those in AD have been demonstrated in brains that have suffered TBI. Studies of traumatized brains a few hours after injury have shown deposition of β-amyloid protein. Moreover, cerebral Aβ deposition was described in both short and long term brain trauma survivors. Production of APP was also noticed to be increased after TBI (Szczygielski et al. 2005). While the underlying biological explanation for the association between TBI and AD is still unknown, similarities in pathological features suggest a possible molecular link.

1.4.5 Vascular Disease

Cerebrovascular disease refers to any disease in which the arteries in the brain or arteries connected to the brain are defective or blocked. Studies have suggested an association between AD and factors that predispose individuals to cerebrovascular disease. Events which restrict blood supply to the brain, such as stroke, have been associated with increasing an individual's risk of AD up to three fold, likely due to increased stress on the brain caused by these events. Factors that predispose individuals to cerebrovascular disease include hypertension, atrial fibrillation, carotid thickening, aortic sclerosis, and diabetes (Kalaria 2003).

1.4.6 Cholesterol Levels

The brain is a cholesterol-rich organ that maintains cholesterol homeostasis separately from the rest of the body, due to the blood-brain barrier (Hartmann et al. 2007). Cholesterol levels in the brain are strictly regulated by local synthesis and elimination of cholesterol in the form of 24(S)-hydroxycholesterol and lipoprotein carrier *APOE*. Lipids may play a role in APP processing,

which leads to production of the 42 amino acid amyloid β peptide (Aβ-42). The action of γ secretase takes place in the lipid membrane, therefore lipid environment of the membrane may play a role in enzyme activity and the development of Aβ and AD (Hartmann et al. 2007).

Studies have demonstrated that cholesterol levels are associated with AD risk. One study of men 70-89 years of age showed that high serum cholesterol level was highly correlated with the prevalence of AD in the study population. Additionally, individuals who take HMG-CoA reductase inhibitors (commonly known as statins), which lower cholesterol levels, have been shown to have lower prevalence and mortality from AD. The exact role of cholesterol in AD pathogenesis is unknown, yet there is evidence that there is a dose-dependent relationship between the level of cholesterol and the average onset of AD (Reiss 2005). Additionally, an increase in cholesterol levels is a risk factor for cerebrovascular disease, which is also a possible risk factor for AD (Kalaria 2003).

1.5 GENETIC SUSCEPTIBILITY

1.5.1 Genetics of Early Onset Alzheimer's Disease

Although AD is most commonly seen in the population age 65 and older, AD accounts for 30% of all nontraumatic dementia in individuals 30-64 years of age (Filley et al. 2007). The three genes known to be associated with early onset AD (EOAD) are *PSEN1*, *PSEN2*, and *APP*. These genes code for presenilin 1 (PSEN1), presenilin 2 (PSEN2) and amyloid precursor protein (APP) (Ertekin-Taner 2007). Mutations in these genes typically follow an autosomal dominant pattern of inheritance. Approximately 30-50% of autosomal dominant familial AD cases have

identifiable mutations in one of these three genes (Filley et al. 2007). Mutations in these three genes are thought to account for 50% of all EOAD cases (Kamboh 2004).

APP is located in chromosome 21q21.2, *PSEN1* at 14q24.3, and *PSEN2* at 1q41 (Ertekin-Taner 2007). Many mutations have been identified and reported in these genes. As of March 2010, 178 mutations in *PSEN1*, 14 mutations in *PSEN2*, and 39 mutations in *APP* have been reported (http://www.molgen.ua.ac.be/Admutations 2010).

The gene products of *PSEN1* and *PSEN2* are both present in brain tissue. *PSEN1* and *PSEN2* are part of the presenilin family, the members of which are involved in various biological processes. Presenilin 1 and presenilin 2 gene products have structural similarities, which suggests that their physiological functions may overlap. It has been hypothesized that *PSEN1* and *PSEN2* code for the active sites of the γ-secretase enzyme. Therefore, mutations in *PSEN1* or *PSEN2* would directly influence APP processing, leading to an increased production of Aβ-42 (Zekanowski et al. 2004).

1.5.2 Genetics of Late Onset Alzheimer's Disease

The etiology of LOAD has been particularly difficult to research due to its complexity. LOAD is thought to be a multifactorial condition in which genes and environment play a role. Liddell et al. (2001) reported a predicted risk of developing AD in first degree relatives of probands with AD as having a 3 to 4-fold increased risk compared to the general population. It is estimated that heritability of LOAD is between 58-79% (Gatz et al. 2006). A study performed by Reynolds et al. (2006) on the heritability of the age of onset of cognitive dysfunction estimated that approximately 35% of variation was due to genetic factors and 65% was attributed to environmental factors.

To date, the only known significant genetic risk factor for LOAD is apolipoprotein E (*APOE),* located on chromosome 19q13 (Borgaonkar et al. 1993). Variations at codons 112 and 158 lead to three common alleles: *APOE*2*, *APOE*3*, and *APOE*4*. The *APOE*2* allele seems to be protective against AD, while the *APOE*4* allele significantly increases an individual's risk. The *APOE*4* allele has a frequency of approximately 14% in the U.S. population. The effect of *APOE*4* on AD risk is dose-dependent. AD risks for *APOE*4* carriers are estimated to be 5 to 15 times greater than that of the general population, depending on whether the individual has one or two copies (Cupples et al. 2004). Additionally, *APOE*4* carriers have an earlier age at onset of AD, which is also dose-dependent (Panza et al. 2004).

APOE is known to be associated with coronary artery disease (CAD) and the development of atherosclerosis, which may contribute to the pathogenesis of AD, considering that vascular disease has been linked to AD. Additionally, *APOE* is believed to be active in the central and peripheral nervous systems. *APOE* is thought to play a role in the mobilization and redistribution of cholesterol and in repair, growth, and maintenance of myelin and neuronal membranes during development and injury. *APOE* occurs in excessive amounts and co-localizes with Aβ plaques in the brains of individuals with AD. *APOE* has also been shown to have a strong stimulatory role in the polymerization of Aβ into amyloid filaments in *in vitro* studies. One hypothesis of *APOE*'s role in the molecular etiology of AD is that the *APOE*4* isoform binds better with Aβ than the *APOE*3* isoform and that *APOE*4* acts as a molecular chaperone that facilitates Aβ accumulation in plaques (Panza et al. 2004).

Although *APOE* is the only confirmed risk factor for LOAD, not all cases carry an *APOE* risk allele, and *APOE* does not explain the entire heritability of LOAD (Ertekin-Taner 2007,

Kamboh 2004). For this reason, researchers continue to search for additional contributing genetic factors in LOAD.

1.5.3 Genome Wide Association Studies

SNP-based genome wide association studies (GWAS) examine the entire human genome for SNP association to specific diseases, rather than examining SNPs that are biological and/or positional candidates. Studies on multifactorial conditions, such as LOAD, can utilize SNP-based GWAS to look for possible genetic factors. Recently several GWAS have been carried out to identify additional genes for AD.

A genome wide association study by Bertram et al. (2008) identified rs179943, rs2049161, rs3826656, and rs11159647 to be associated with AD risk. SNP rs17993 is located in the ataxin-1 gene (*ATXN1*), at 6p23. The *ATXN1* gene product, ataxin-1 protein, has been demonstrated to be present in various brain regions and in nonneuronal tissue, such as heart, skeletal muscle, and liver. ATXN1 contains a CAG repeat region which codes for a polyglutamine tract which appears to have no effect on protein transcription or translation (Servadio et al. 1995). One group demonstrated that the mutated (expanded CAG tract) form of the protein appears to be more resistant to degradation (Cummings et al. 1999). Studies have also suggested that the resistance to degredation may lead to accumulation of the mutant protein, noting that the pathologic hallmark of neurodegenerative disorders associated with expanded polyglutamine repeats is the accumulation of polyglutamine inclusions. SNP rs2049161 is located at 18p11.3, approximately 150 kb from the psuedogene glyceraldehyde-3-phosphate dehydrogenase-like 11 (*GAPDHL11*). This pseudogene bears resemblance to the glyceraldehyde-3-phosphate dehydrogenase, which is an enzyme that functions as part of the glycolysis pathway.

Evidence has shown that GAPDH glycolytic function is impaired in Alzheimer's disease despite unchanged gene expression (Mazzola and Sirover 2001). SNP rs3826656 is located in the myeloid differentiation antigen CD33 gene (*CD33*), which located between 19q13.3 and 19q13.4. The *CD33* gene product has been demonstrated to be expressed by human monocytes, promyelocytes, myeloid blasts, some acute undifferentiated leukemias, and occasionally by acute lymphoblastic leukemias (Peiper et al. 1988). SNP rs11159647 is located at 14q31, approximately 100kb from the hypothetical gene LOC100287804.

A genome wide association study by Beecham et al. (2008) identified rs7893928, rs16934131, rs7894737, rs11244841, rs3746319, rs11610206 to be associated with AD risk. SNP rs7893928 is located at 10q11.21, approximately 175 kb from pseudogene ribosomal protein L9 pseudogene 21 (RPL921). SNP rs16934131 is located in the potassium channel, calcium-activated, large conductance, subfamily M, alpha member 1 gene (*KCNMA1*), at 10q22.3. *KCNMA1* has been suggested to be associated with innate immunity, cochlear hearing, vascular dysfunction, epilepsy, and mental retardation. Knockout mice have been engineered by targeted disruption of the *KCNMA1* gene which showed obvious ataxia, moderate vascular dysfunction, and independent neurologic deficits, including cerebellar dysfunction in the form of abnormal eye-blink reflexes, abnormal motion, and deficiency in motor coordination (Sausbier et al. 2004). SNP rs7894737 is located in the SorCS receptor 3 gene (*SORCS3*), at 10q23.3. *SORCS1* and *SORCS2* were identified as orthologs of VPS10 domain-containing receptor genes in mice. *SORCS3* was named for similarities to the first two genes, but its exact function remains unknown (Hampe et al. 2001, Hermey et al. 1999). SNP rs11244841 is located in a disintigrin and metalloproteinase domain 12 (*ADAM12*), which is located at 10q26.3. The function of *ADAM12* remains largely unknown, however studies have shown that *ADAM12* expression

increases during muscle regeneration in mice (Galliano et al. 2000). SNP rs3746319 is located in the zinc finger protein 224 gene (*ZNF224*), between 19q13.2 and 19qter. The *ZNF224* gene has been shown to have wide expression that was highest in hematopoietic tissues and cell lines (Han et al. 1999). The *ZNF224* gene product has been suggested to function as a repressor, inhibiting transcription of various genes, including Wilms tumor 1 (*WT1*) (Lee et al. 2002). SNP rs11610206 is located in the family with sequence similarity 113, member B gene (*FAM113B*), at 12q13.

A genome wide association study by Reiman et al. (2007) identified rs2373115 to be a modifier of AD risk in *APOE4* carriers. This SNP is located in the *GRB*-associated binding protein 2 gene *(GAB2),* between 11q13.4 and 11q13.5. *GAB2* has been implicated in coupling cytoplasmic-nuclear signal transduction (Zhao et al. 1999). GAB2 is located in a region commonly amplified in human breast cancer, and is overexpressed in breast cancer cell lines and primary tumors (Bentires-Alj et al. 2006).

A genome wide association study by Carrasquillo et al. (2009) identified rs5984894 to be associated with AD risk. SNP rs5984894 is located in the protocadherin 11, X-linked gene (*PCDH11X*), at Xq21.3. *PCDH11X* is paired with *PCDH11Y* in a human-specific X/Y gene pair located in the Xq21.2/Yp11.2 homology block. They are both members of a family of calcium-dependent cell adhesion and recognition proteins and are particularly prevalent in the central nervous system (Yoshida and Sugano 1999).

A recent GWAS by Harold et al. (2009) of over 16,000 individuals revealed significant associations between AD and SNPs at two loci, 5’ to the *PICALM* gene and the *CLU* gene (also known as *APOJ*). Additionally, a GWAS by Lambert et al. (2009) identified two loci, *CLU* and *CR1*, which appeared to be associated with AD. *PICALM* encodes the phosphatidylinositol-

binding clathrin assembly protein at 11q14. *CLU* encodes for clusterin or apolipoprotein J at 8p21-p12. *CR1* encodes the complement component receptor 1 at 1q32. Roles of *CLU* and *CR1* in the clearance of Aβ have been suggested in previous biological studies (Lambert et al. 2009). Our group has examined *APOJ, CR1* and *PICALM* in a large case-control study comprised of 2,707 individuals. No association was observed with *CR1* SNPs, however, a trend of association was seen with *PICALM* and *CLU* SNPs (Kamboh et al. 2010).

In this study, a total of 12 SNPs were chosen from studies performed by Bertram et al. (2008), Beecham et al. (2009), Reiman et al. (2007), and Carrasquillo et al. (2009) for a replication study to determine if the associations seen between SNPs and AD risk in the respective studies were present in our population.

1.5.4 SNPs in Putative Imprinted Genes

Positional candidate SNPs were chosen on chromosome 10 because of observed linkage peaks which indicate that these regions are segregating with risk for AD (Kamboh 2004). The parent of origin effect has been hypothesized to play a role in AD risk (Bassett et al. 2002). One of the major mechanisms of parent of origin is imprinting. One study by Luedi et al. (2007) predicted potential imprinted genes in the human genome. Out of the SNPs available through TaqMan assays, we chose 10 SNPs that would best cover the hypothesized imprinted genes within the linkage region of chromosome 10.

1.6 STUDY OBJECTIVES

The objective of this thesis project is to try to replicate the associations between AD and SNPs found by Bertram et al. (2008), Beecham et al. (2009), Reiman et al. (2007), and Carrasquillo et al. (2009) and test for any possible associations between SNPs in putative imprinted genes on chromosome 10 and AD risk. Additionally, we will test if any of these SNPs have associations with age at onset (AAO), disease duration, and mini mental state examination (MMSE) score.

2.0 METHODS

2.1 SUBJECTS

2.1.1 Sample Population

Subjects were 1,969 Caucasian American individuals, including 993 LOAD cases and 976 older controls. In the AD group, 67.1% were female and 8.3% were autopsy-confirmed. The mean age-at-onset in AD cases was 72.9 ± (SD) 6.3 years. In the control group, 60.2% were female and 1.2% were autopsy confirmed. The mean age of controls at baseline was 74.2 ± 6.2 years. The subjects were recruited with informed consent and the study was approved by the University of Pittsburgh Institutional Review Board

DNA was isolated from blood using the QIAamp blood DNA Maxi kit protocol (Qiagen, Valencia, CA), and from brain tissue using the QIAamp DNA Mini kit protocol. A small number of samples with insufficient amounts of DNA were amplified using the GenomiPhi kit (GE Healthcare).

2.2 GENOTYPE DETERMINATION

Information on SNPs studied in this thesis project can be found in Table 1.

Table 1. Study SNPs

dbSNP	Chromosome Location	Mapped ID	Rationale for Examination
rs179943	6p23	*ATXN1*	GWAS
rs35011589	10p14	*GATA3*	Putative Imprinted Gene
rs7893928	10q11.21	200kb 5' of *CXCL12*	GWAS
rs16934131	10q22.3	*KCNMA1* Intron	GWAS
rs7894737	10q25.1	*SORCS3* Intron	GWAS
rs11244841	10q26.3	*ADAM12* Intron	GWAS
rs11542965	10q26.3	*PAOX*	Putative Imprinted Gene
rs1049951	10q26.3	*PAOX*	Putative Imprinted Gene
rs11101654	10q26.3	*VENTX2*	Putative Imprinted Gene
rs11599284	10q26.3	*VENTX2*	Putative Imprinted Gene
rs2270192	10q26.3	*VENTX2*	Putative Imprinted Gene
rs2240891	10q26.3	*VENTX2*	Putative Imprinted Gene
rs7907431	10q26.3	*C10orf93*	Putative Imprinted Gene
rs4500406	10q26.3	*C10orf93*	Putative Imprinted Gene
rs12781609	10q26.3	*C10orf93*	Putative Imprinted Gene
rs2373115	11q14.1	*GAB2*	GWAS
rs11610206	12q13	*FAM113B*	GWAS
rs11159647	14q31	100kb from *LOC100287804*	GWAS
rs2049161	18p11.3	150kb 5' of *GAPDHL11*	GWAS
rs3746319	19q13.2	*ZNF224*	GWAS
rs3826656	19q13.3	1.5 kb 5' of *CD33*	GWAS
rs5984894	Xq21.3	*PCDH11X* Intron	GWAS

2.2.1 TaqMan Assay

The genotypes for 19 SNPs were determined using TaqMan SNP genotyping Assays (Applied Biosystems, Foster City, CA). The SNP number and assay identification numbers for each SNP are listed in Table 2.

Table 2. TaqMan SNP genotyping Assays

dbSNP	Chromosome Location	Gene	Assay ID	Sequence [VIC/FAM]	dbSNP alleles
rs179943	6p23	*ATXN1*	C__10031092_10	A/G	C/T
rs35011589	10p14	*GATA3*	C__25760909_10	C/G	C/G
rs7893928	10q11.2	*CXCL12*	C__26949936_10	C/T	C/T
rs16934131	10q22.3	*KCNMA1*	C___2003447_10	C/T	C/T
rs7894737	10q25.1	*SORCS3*	C__11759269_10	A/C	A/C
rs11244841	10q26.3	*ADAM12*	C___1701036_10	C/T	C/T
rs11542965	10q26.3	*PAOX*	C__61810215_10	C/T	G/A
rs1049951	10q26.3	*PAOX*	C___7468389_10	G/A	C/T
rs11101654	10q26.3	*VENTX2*	C__31984593_10	A/C	A/C
rs11599284	10q26.3	*VENTX2*	C__31984728_10	A/G	A/G
rs2270192	10q26.3	*VENTX2*	C__27105356_10	A/G	C/T
rs2240891	10q26.3	*VENTX2*	C__15872002_10	G/A	C/T
rs7907431	10q26.3	*C10orf93*	C__29348997_10	T/C	C/T
rs4500406	10q26.3	*C10orf93*	C__27905453_10	G/C	C/G
rs12781609	10q26.3	*C10orf93*	C____204339_10	T/C	C/T
rs2049161	18p11.3	*GAPDHL11*	C__11499278_10	A/C	A/C
rs3746319	19q13.2	*ZNF224*	C__22274939_10	A/G	C/T
rs3826656	19q13.3	*CD33*	C___1487396_10	A/G	A/G
rs5984894	Xq21.3	*PCDH11X*	rs5984894_M	T/C	A/G

The TaqMan analysis was performed on samples placed on 384-well plates. Each plate had a mixture of cases and controls, with ten percent of all samples repeated in order to assess error rate. Reaction components and cycling conditions were followed according to Applied BioSystems TaqMan protocol.

The biochemical principle of TaqMan is briefly outlined as follows. The DNA is first amplified by utilizing the polymerase chain reaction (PCR) using specific probe-labeled primers, then the samples are read using the 7900HT Real-Time PCR system machine (Applied Biosystems). The samples are first placed on 384 well plates. A reaction mixture is then added to each sample which contains master mix, genotyping assay, and water. The genotyping assay

consists of sequence specific forward and reverse primers, one probe labeled at the 5’ end with VIC dye and one probe labeled at the 5’ end with FAM dye. A nonfluorescent quencher is attached to the 3’ end of each of these probes. The probes lie between the two primers and bind specifically to the SNP of interest.

The 5’ VIC and FAM probes that are contained in the reaction mixture contain a “reporter” at one end and a “quencher” at the other end. When the reporter and quencher remain together, there is no fluorescence. Yet when they are separated, the reporter is no longer “quenched” and it fluoresces. These probes bind to the region containing the SNP of interest, each binding specifically to different alleles. AmpliTaq Gold® Polymerase with 5’ exo-nuclease activity is used in the PCR reaction. During the annealing process, the polymerase binds to the primers and begins elongation of the new DNA strand. As it proceeds along the DNA strand, the exo-nuclease activity of the polymerase removes any double-stranded DNA (dsDNA) that may be impeding its path. Therefore, when the polymerase reaches the attached probe during elongation, the 5’ exo-nuclease activity cleaves the probe. The probes that are cleaved allow for the VIC/FAM dye to separate from the quencher, which allows for the dye to fluoresce. The reaction is illustrated in Figure 1.

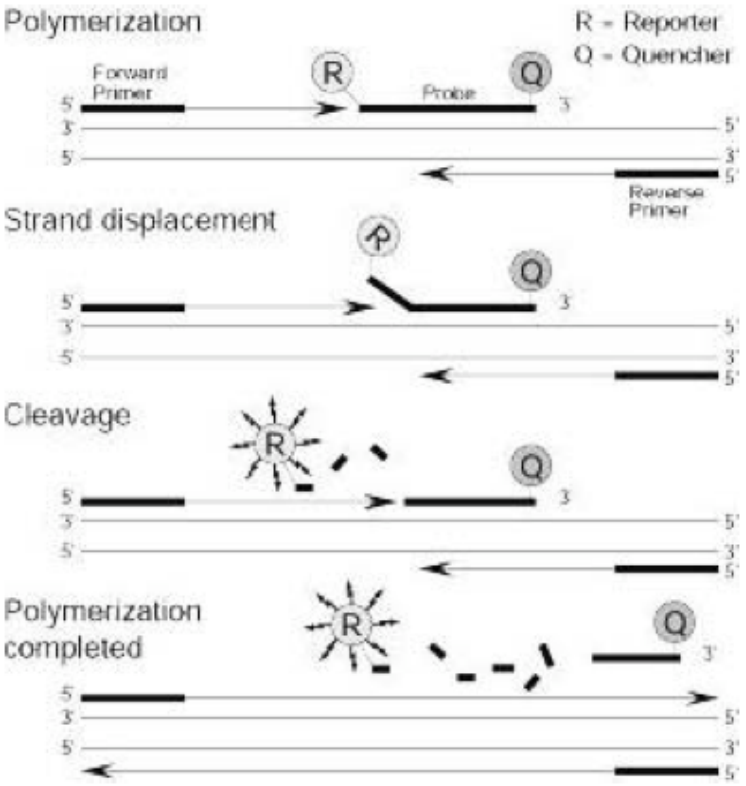

Figure 1. TaqMan Chemical Reaction

The product is analyzed using the 7900HT Fast Real-Time PCR system. The fluorescence that is released by the probes is specific to VIC or FAM, and therefore each corresponding allele, and can be detected by the 7900HT Fast Real-Time PCR system machine. The fluorescence of each product is measured and placed on a graph accordingly. The output allows us to distinguish between the different clusters of fluorescence intesnsity that are seen for each genotype. An example of the results from a TaqMan analysis can be seen in Figure 2.

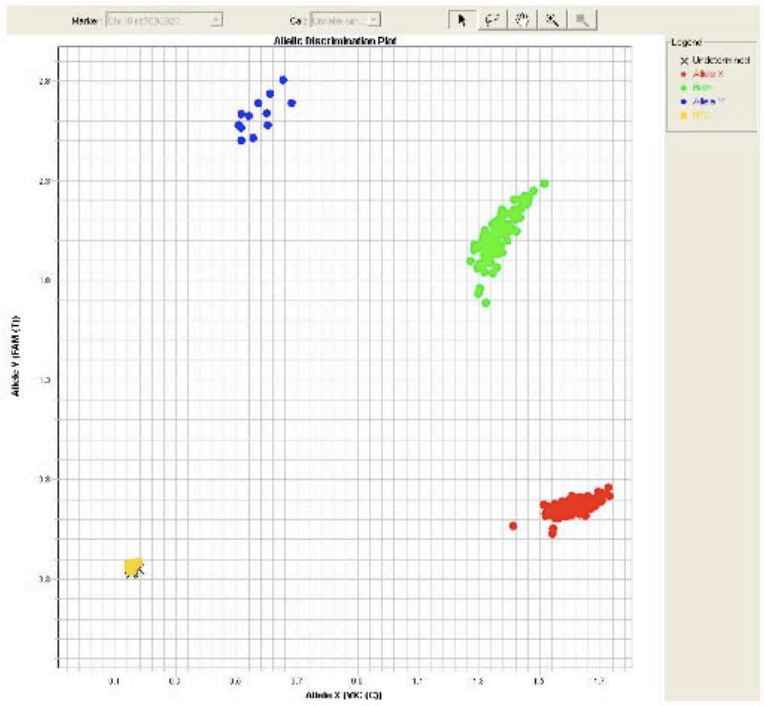

Figure 2. Example Result from TaqMan Analysis

2.2.2 Pyrosequencing

Pyrosequencing™ (PSQ) was developed by Ronaghi et al. (1998) and is based on real time pyrophosphate (Ppi) detection, which is released as a result of nucleotide incorporation by DNA polymerase. Pyrosequencing can be used to determine sequences of short segments of DNA and therefore can be utilized in the genotype analysis of SNP.

In short, 40 µL of binding buffer (10mM Tris-HCl pH 7.6, 2M NaCl, 1mM EDTA, 0.1% Tween20), 2µL Streptavidin Sepharose beads (Amersham Biosciences, Piscataway, NJ) and 18 µL of high purity water were added to 20µL of the amplified biotinylated PCR product. This mix was then shaken for 10 minutes, which allowed binding of the biotinylated PCR DNA to the streptavidin beads. DNA molecules, now immobilized on the streptavidin beads, were next captured onto filter probes using a vacuum prep tool. The beads were then washed in 70%

ethanol for 10 seconds, denaturation solution (0.2M NaOH) for 10 seconds, and in a wash buffer (10mM Tris Acetate pH7.6) for 20 seconds. The purpose of these steps are to watch and denature the DNA so that the double stranded DNA (dsDNA) separates into its single stranded DNA (ssDNA) templates. Next, the suction was stopped and the ssDNA (on beads) were released into the wells of a PSQ 96 plate containing 9.5µL of annealing buffer and 0.5µL of sequencing primer in each well. This 96 well plate was next placed on a 90ºC heating block with heated lid for 2 minutes, followed by 2 minutes at room temperature. This heating process promotes annealing of the ssDNA with the sequencing primer. Analysis begins with the dispensation of the enzyme and substrate into each well of the plate. Essentially, DNA polymerase catalyzes synthesis of a DNA strand complimentary to the original sequence, which is coupled to a chemiluminescent reaction. This generates light proportional to the number of nucleotides incorporated into the elongating strand. The light is detected by a charge coupled device (CCD) camera that is located inside the PSQ system. PSQ software automatically analyzes the quantitative data and presents the data in the form of programs, which can easily be evaluated by the user. Examples of pyrograms for various genotypes can be seen in Figures 3 - 5. Figures 6 and 7 show an outline view of Pyrosequencing™ assay and the general biochemical principles behind the Pyrosequencing™ reaction system.

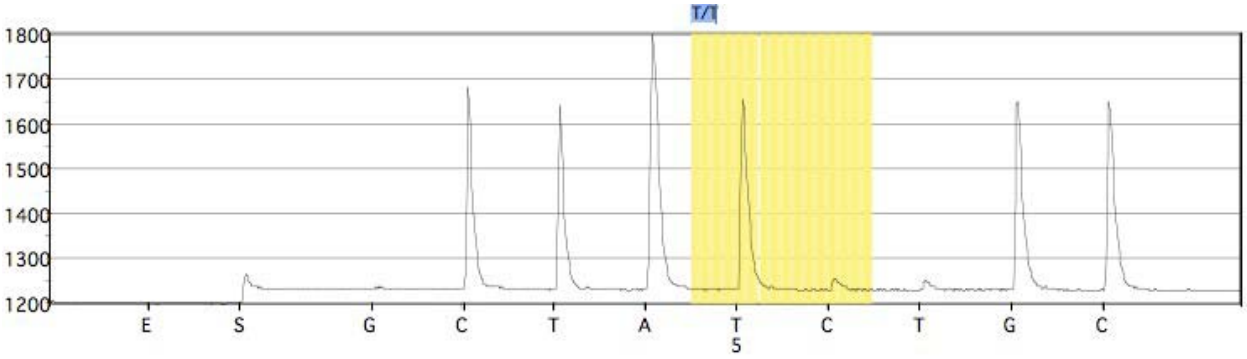

Figure 3. rs11610206 pyrogram example of TT genotype

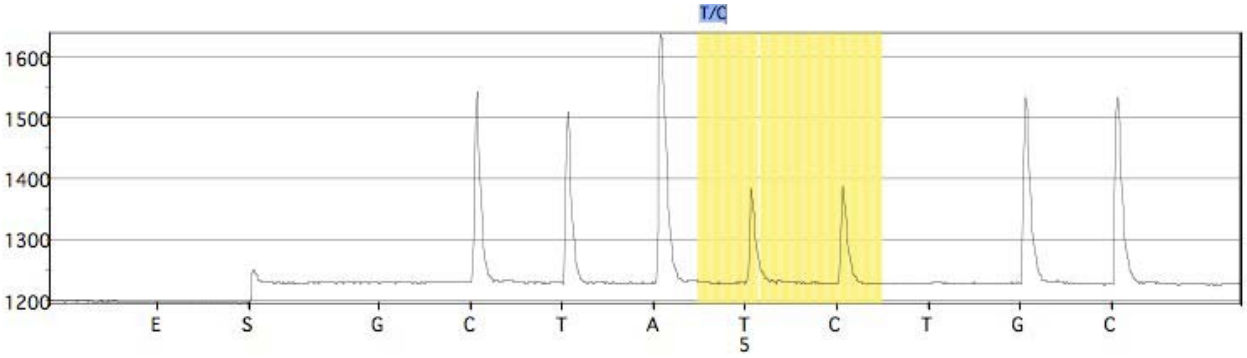

Figure 4. rs11610206 pyrogram example of TC genotype

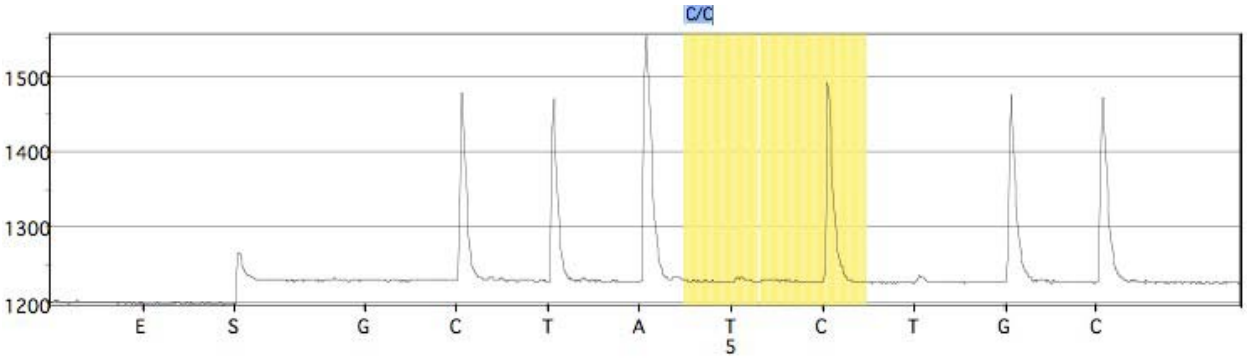

Figure 5. rs11610206 pyrogram example of CC genotype

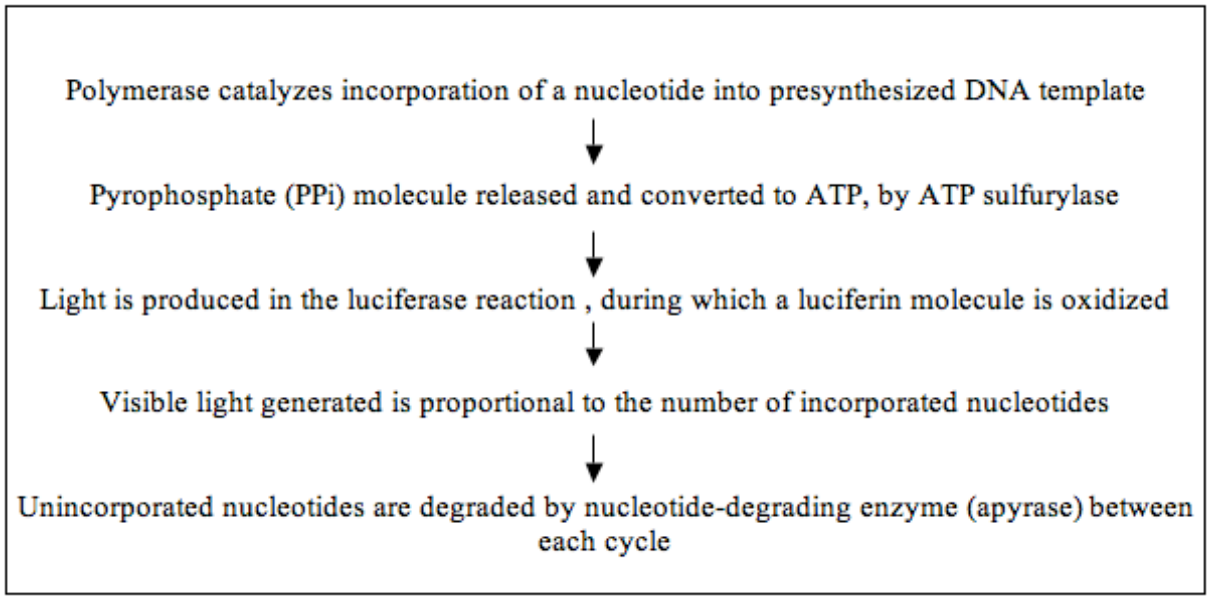

Figure 6. An outline view of Pyrosequencing™ assay

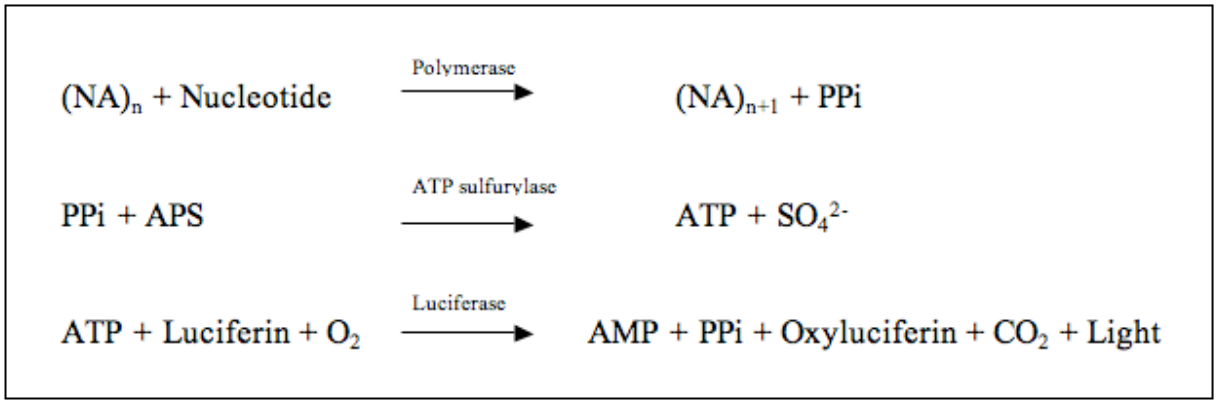

Figure 7. The general principal behind Pyrosequencing™ assay (adapted from Ronaghi et al. 2001)

The genotypes for three SNPs (rs2373115, rs11610206, and rs11159647) were determined by Pyrosequencing™ on the PSQ™ HS 96 system (Qiagen, Valencia, CA). Table 3 contains the sequences of the primers used and the target sequence to analyze for each SNP genotyped by Pyrosequencing™. The Pyrosequencing™ analysis was performed on samples that were placed in 96 well plates. Approximately 10% of samples were repeated in order to assess error rate.

Table 3. Primer sequences and sequences analyzed for SNPs genotyped by Pyrosequencing

rs2373115 *GAB2*	Sequence to Analyze	AG/TGCACTC
	Sequence Primer	GACATGGATTTATAGTCCG
	Forward Primer	AGACTTATGCGGACATGGATTT
	Reverse (biotinylated) Primer	AGAAATGTTCGACCATCAATACTC
rs11610206 *FAM113B*	Sequence to Analyze	GCTAT/CTGC
	Sequence Primer	GGTGATCTTTAGCCTCC
	Forward Primer	GTTAATCTGCCTTGTGTCAATTTG
	Reverse (biotinylated) Primer	ACTGGCACAATGGGCTAGGT
rs11159647 *LOC10028 7804*	Sequence to Analyze	GCATG/AGTA
	Sequence Primer	CATATAGCAAAGCTGCA
	Forward Primer	GATGAGCTGGAGCCCATAAT
	Reverse (biotinylated) Primer	AGGGCCCACACTAAGTACTCATCT

2.3 STATISTICAL METHODS

Allele and genotype frequencies were calculated by the direct allele-counting method. Goodness of fit to Hardy-Weinberg Equilibrium was tested using X^2 test. Differences between genotype and allele frequencies in cases and controls were tested with the X^2 test. The X^2 test was also used to test for differences between cases and controls, stratified by *APOE*4* status and stratified by sex. These statistics were calculated using PLINK v1.07 (Purcell et al. 2007).

Logistic regression was used to calculate odds ratios and confidence intervals, adjusted by age (in controls) or AAO (in cases) and *APOE*4* status. Linear regression analysis was performed for AAO, disease duration (age at death – age at diagnosis), and baseline MMSE score. We used AAO and sex as covariates for disease duration, age at testing as a covariate for MMSE score, and *APOE*4* status as a covariate for AAO. These calculations were done using PLINK v1.07 (Purcell et al. 2007).

Kaplan-Meier analysis for AAO, in all subjects and stratified by *APOE*4* status, and regression analysis for AAO were performed using R v2.7.2 (R Development Core Team 2007). Power was estimated using R v2.7.2 (R Development Core Team 2007).

3.0 RESULTS

3.1 APOE POLYMORPHISM IN AD CASES AND CONTROLS

APOE genotype frequencies are presented in Table 4. In cases, the 3/4 genotype was the most frequent genotype, representing 46% of the total case genotypes. The 3/3 genotype was the second most frequent, making up 38% of the total case genotypes. The remainder of the genotypes were below 10%. In controls, the 3/3 genotype was the most frequent, making up 65% of the total control genotypes. The remainder of the genotypes of controls were below 20%. While the frequency of the *APOE*4* allele was significantly higher in cases than controls ($p<0.001$), the frequency of the *APOE*2* allele was lower in cases than controls ($p<0.001$).

Table 4. APOE polymorphism AD cases and controls

	AD Cases		**Controls**	
	n	**freq**	***n***	**freq**
3/3	394	0.384	654	0.653
3/4	470	0.459	165	0.165
2/3	39	0.038	141	0.141
2/4	28	0.027	22	0.022
4/4	92	0.090	14	0.014
2/2	2	0.002	5	0.005
Total (n)	1025		1001	
*APOE*3*	1297	0.633	1614	0.806
*APOE*4*	682	0.333	215	0.107
*APOE*2*	71	0.035	173	0.086

3.2 GENOTYPING ERROR RATE

Of the 22 SNPs examined, the estimated error rate for rs2373115, rs11610206, rs11159647 (genotyped by Pyrosequencing) were 2.3%, 1.7%, and 3.3%, respectively. The genotyping error rate for all remaining SNPs using TaqMan was estimated to be <0.6%.

3.3 MONOALLELIC SNPS

Six of the 22 studied SNPs were found to be monoallelic when screened in up to 1978 individuals. These SNPs include rs35011589, rs11542965, rs2270192, rs11101654, rs4500406, and rs7907431. These SNPs are excluded from subsequent tables and discussions. Additionally, one SNP, rs2240891, had minor allele frequency (MAF) of 0.002 in 1962 individuals. The minor allele frequency was identical in cases and controls, therefore this SNP is also excluded from subsequent tables and discussions. These seven SNPs were part of the putative imprinted genes on chromosome 10. We did not have prior knowledge about their being monoallelic.

3.4 HARDY WEINBERG EQUILIBRIUM

P-values based on Hardy Weinberg Equilibrium (HWE) analysis are listed in Table 5. All but one of the studied SNPs were found to be in HWE for both cases and controls. One SNP, rs2049161, was found to be out of HWE in the control population. However, when taking the number of tests performed into account, this is not statistically significant.

Table 5. p-values based on HWE for cases and controls

SNP	**Gene**	***p*-value**	
		cases	controls
rs179943	*ATXN1*	0.129	1.000
rs7893928	*CXCL12*	0.803	0.061
rs16934131	*KCNMA1*	0.399	0.617
rs7894737	*SORCS*	0.947	0.501
rs11244841	*ADAM12*	0.804	0.932
rs1049951	*PAOX*	0.058	1.000
rs11599284	*VENTX2*	0.670	0.341
rs12781609	*C10orf93*	0.415	0.057
rs2373115	*GAB2*	0.130	0.829
rs11610206	*FAM113B*	0.123	1.000
rs11159647	*LOC100287804*	0.394	0.550
rs2049161	*GAPDHL11*	0.097	0.002
rs3746319	*ZNF224*	0.915	0.526
rs3826656	*CD33*	0.138	1.000
rs5984894	*PCDH11X*	0.101	0.283

3.5 DISTRIBUTION OF SNPS IN CASES AND CONTROLS

Allele and genotype frequencies for 15 SNPs for cases and controls are given in Table 6. None of the SNPs revealed significant association with AD risk.

Table 6. Distribution of SNPs in cases and controls

SNP/Gene		***n***	**Genotype Counts (frequency)**						***p***	**Allele frequency**		***p***
			CC		CT		TT			C	T	
rs179943	AD	907	784	(0.864)	115	(0.127)	8	(0.009)	0.192	0.928	0.072	0.177
ATXN1	C	914	766	(0.838)	142	(0.155)	6	(0.007)		0.916	0.084	

			CC		CT		TT			C	T	
rs7893928	AD	989	716	(0.724)	250	(0.253)	23	(0.023)	0.523	0.850	0.150	0.854
CXCL12	C	972	707	(0.727)	235	(0.242)	30	(0.031)		0.848	0.152	
			TT		TC		CC			T	C	
rs16934131	AD	993	563	(0.567)	363	(0.366)	67	(0.067)	0.494	0.750	0.250	0.518
KCNMA1	C	976	532	(0.545)	382	(0.391)	62	(0.064)		0.741	0.259	
			AA		AC		CC			A	C	
rs7894737	AD	979	353	(0.361)	469	(0.479)	157	(0.160)	0.724	0.600	0.400	0.500
SORCS	C	976	369	(0.378)	454	(0.465)	153	(0.157)		0.611	0.389	
			TT		TC		CC			T	C	
rs11244841	AD	980	533	(0.544)	382	(0.390)	65	(0.066)	0.687	0.739	0.261	0.391
ADAM12	C	965	543	(0.563)	363	(0.376)	59	(0.061)		0.751	0.249	
			AA		AG		GG			A	G	
rs1049951	AD	991	871	(0.879)	112	(0.113)	8	(0.008)	0.064	0.935	0.065	0.058
PAOX	C	971	875	(0.901)	94	(0.097)	2	(0.002)		0.950	0.050	
			GG		GA		AA			G	A	
rs11599284	AD	992	755	(0.761)	219	(0.221)	18	(0.018)	0.382	0.872	0.129	0.100
VENTX2	C	973	766	(0.787)	191	(0.196)	16	(0.016)		0.885	0.115	
			CC		CT		TT			C	T	
rs12781609	AD	982	388	(0.395)	449	(0.457)	145	(0.148)	0.739	0.624	0.376	0.920
C10orf93	C	974	391	(0.401)	430	(0.441)	153	(0.157)		0.622	0.378	
			GG		GT		TT			G	T	
rs2373115	AD	923	603	(0.653)	295	(0.320)	25	(0.027)	0.650	0.813	0.187	0.899
GAB2	C	915	601	(0.657)	283	(0.309)	31	(0.034)		0.812	0.189	
			TT		TC		CC			T	C	
rs11610206	AD	957	793	(0.829)	152	(0.159)	12	(0.013)	0.554	0.908	0.092	0.853
FAM113B	C	909	752	(0.827)	150	(0.165)	7	(0.008)		0.910	0.090	
			GG		GA		AA			G	A	
rs11159647	AD	942	270	(0.287)	481	(0.511)	191	(0.203)	0.560	0.542	0.458	0.774
LOC100287804	C	913	268	(0.294)	445	(0.487)	200	(0.219)		0.537	0.463	
			AA		AC		CC			A	C	
rs2049161	AD	777	522	(0.672)	221	(0.284)	34	(0.044)	0.376	0.814	0.186	0.301
GAPDHL11	C	840	589	(0.701)	213	(0.254)	38	(0.045)		0.828	0.172	
			GG		GA		AA			G	A	
rs3746319	AD	980	655	(0.668)	292	(0.298)	33	(0.034)	0.858	0.817	0.183	0.705
ZNF224	C	966	641	(0.664)	288	(0.298)	37	(0.038)		0.813	0.187	
			AA		AG		GG			A	G	
rs3826656	AD	811	487	(0.600)	273	(0.337)	51	(0.063)	0.416	0.769	0.231	0.522
CD33	C	875	504	(0.576)	321	(0.367)	50	(0.057)		0.759	0.241	
			GG		GC		CC			T	C	
rs5984894	AD	657	187	(0.285)	307	(0.467)	163	(0.248)	0.134	0.514	0.486	0.996
PCDH11X	C	588	147	(0.250)	308	(0.524)	133	(0.226)		0.514	0.486	

3.6 POWER CALCULATION AT α = 0.05 FOR EXAMINED SNPS

80% power at α = 0.05 to detect the effect of a SNP using an allele based test given the respective sample sizes on disease risk, are shown in Table 7. As indicated in the table, we had sufficient power to detect odds ratios greater than 1.2-1.59 and less than 0.68-0.83.

Table 7. 80% power calculation at α = 0.05 for 15 SNPs

SNP	Gene	*n* cases	*n* controls	> OR for risk allele	< OR for protective allele
rs179943	*ATXN1*	907	914	1.44	0.73
rs7893928	*CXCL12*	989	972	1.31	0.78
rs16934131	*KCNMA1*	993	976	1.24	0.81
rs7894737	*SORCS*	979	976	1.21	0.83
rs11244841	*ADAM12*	980	965	1.24	0.81
rs1049951	*PAOX*	991	971	1.59	0.68
rs11599284	*VENTX2*	992	973	1.36	0.76
rs12781609	*C10orf93*	982	974	1.21	0.83
rs2373115	*GAB2*	923	915	1.29	0.79
rs11610206	*FAM113B*	957	909	1.43	0.73
rs11159647	*LOC100287804*	942	913	1.21	0.83
rs2049161	*GAPDHL11*	777	840	1.31	0.78
rs3746319	*ZNF224*	980	966	1.28	0.79
rs3826656	*CD33*	811	875	1.26	0.80
rs5984894	*PCDH11X*	657	588	1.20	0.83

3.7 ADJUSTED ODDS RATIOS

Age/AAO-, sex- and *APOE*4*-adjusted odds ratios and confidence intervals are given in Table 8. Odds ratios for covariates were similar across all SNPs. No statistically significant associations were found.

Table 8. Odds ratios and confidence intervals for 15 SNPs , adjusted for age/AAO and *APOE*4* status

SNP	Gene	OR (95% CI)	*p*-value
rs179943	*ATXN1*	0.85 (0.65-1.10)	0.2218
rs7893928	*CXCL12*	1.04 (0.86-1.26)	0.6620
rs16934131	*KCNMA1*	0.89 (0.77-1.05)	0.1815
rs7894737	*SORCS*	0.99 (0.87-1.15)	0.9615
rs11244841	*ADAM12*	1.09 (0.93-1.28)	0.3016
rs1049951	*PAOX*	1.32 (0.98-1.76)	0.0644
rs11599284	*VENTX2*	1.07 (0.87-1.32)	0.5304
rs12781609	*C10orf93*	1.01 (0.88-1.16)	0.8610
rs2373115	*GAB2*	0.99 (0.83-1.20)	0.9900
rs11610206	*FAM113B*	1.06 (0.83-1.35)	0.6264
rs11159647	*LOC100287804*	0.99 (0.87-1.15)	0.9814
rs2049161	*GAPDHL11*	1.11 (0.92-1.34)	0.2786
rs3746319	*ZNF224*	0.94 (0.79-1.12)	0.4785
rs3826656	*CD33*	0.90 (0.76-1.07)	0.2513
rs5984894	*PCDH11X*	0.98 (0.84-1.14)	0.7945

-

3.8 STRATIFICATION BY *APOE*

APOE stratified data for 15 SNPs are presented in Table 9. No statistically significant differences were found in genotype or allele frequencies between cases and controls, regardless of *APOE*4* status.

Table 9. Distribution of 15 SNPs, based on *APOE*4* carriers and non-*APOE*4* carriers

SNP/Gene				Genotype Count (allele frequency)						*p*	Allele frequency		*p*
			n	CC		CT		TT			C	T	
rs179943 *ATXN1*	No E4	AD	393	341	(0.868)	47	(0.120)	5	(0.013)	0.092	0.927	0.073	0.303
		C	730	610	(0.836)	116	(0.159)	4	(0.005)		0.915	0.085	
	E4	AD	514	443	(0.862)	68	(0.132)	3	(0.006)	0.628	0.928	0.072	0.533
		C	183	155	(0.847)	26	(0.142)	2	(0.011)		0.918	0.082	
			n	CC		CT		TT			C	T	
rs7893928 *CXCL12*	No E4	AD	417	297	(0.712)	110	(0.264)	10	(0.024)	0.713	0.844	0.156	0.701
		C	772	547	(0.709)	200	(0.259)	25	(0.032)		0.838	0.162	
	E4	AD	572	419	(0.733)	140	(0.245)	13	(0.023)	0.153	0.855	0.145	0.133
		C	192	153	(0.797)	34	(0.177)	5	(0.026)		0.885	0.115	
			n	TT		TC		CC			C	T	
rs16934131 *KCNMA1*	No E4	AD	423	256	(0.605)	142	(0.336)	25	(0.059)	0.184	0.773	0.227	0.095
		C	776	427	(0.550)	298	(0.384)	51	(0.066)		0.742	0.258	
	E4	AD	570	307	(0.539)	221	(0.388)	42	(0.074)	0.637	0.733	0.268	0.941
		C	192	101	(0.526)	80	(0.417)	11	(0.057)		0.734	0.266	
			n	AA		AC		CC			A	C	
rs7894737 *SORCS*	No E4	AD	413	149	(0.361)	206	(0.499)	58	(0.140)	0.525	0.610	0.390	0.617
		C	775	301	(0.388)	360	(0.465)	114	(0.147)		0.621	0.379	
	E4	AD	566	204	(0.360)	263	(0.465)	99	(0.175)	0.863	0.593	0.407	0.594
		C	194	67	(0.345)	90	(0.464)	37	(0.191)		0.577	0.423	
			n	TT		TC		CC			T	C	
rs11244841 *ADAM12*	No E4	AD	412	220	(0.534)	167	(0.405)	25	(0.061)	0.713	0.737	0.263	0.643
		C	764	424	(0.555)	291	(0.381)	49	(0.064)		0.745	0.255	
	E4	AD	568	313	(0.551)	215	(0.379)	40	(0.070)	0.560	0.740	0.260	0.300
		C	193	113	(0.585)	70	(0.363)	10	(0.052)		0.767	0.233	
			n	AA		AG		GG			A	G	
rs1049951 *PAOX*	No E4	AD	421	366	(0.869)	52	(0.124)	3	(0.007)	0.151	0.931	0.069	0.061
		C	775	699	(0.902)	74	(0.095)	2	(0.003)		0.950	0.050	
	E4	AD	570	505	(0.886)	60	(0.105)	5	(0.009)	0.013	0.939	0.061	0.463
		C	195	175	(0.897)	20	(0.103)	0	(0.000)		0.949	0.051	
			n	GG		GA		AA			G	A	
rs11599284 *VENTX2*	No E4	AD	421	334	(0.793)	78	(0.185)	9	(0.021)	0.881	0.886	0.114	0.969
		C	777	613	(0.789)	150	(0.193)	14	(0.018)		0.886	0.115	
	E4	AD	571	421	(0.737)	141	(0.247)	9	(0.016)	0.322	0.861	0.139	0.231
		C	195	152	(0.779)	41	(0.210)	2	(0.010)		0.885	0.115	
			n	CC		CT		TT			C	T	
rs12781609 *C10orf93*	No E4	AD	419	154	(0.368)	196	(0.468)	69	(0.165)	0.501	0.601	0.399	0.292
		C	778	313	(0.402)	344	(0.442)	121	(0.156)		0.623	0.377	
	E4	AD	563	234	(0.416)	253	(0.449)	76	(0.135)	0.618	0.640	0.360	0.416
		C	196	78	(0.398)	86	(0.439)	32	(0.163)		0.617	0.383	
			n	GG		GT		TT			G	T	
rs2373115 *GAB2*	No E4	AD	383	242	(0.632)	127	(0.332)	14	(0.037)	0.564	0.798	0.202	0.328
		C	729	484	(0.664)	220	(0.302)	25	(0.034)		0.815	0.185	
	E4	AD	540	361	(0.669)	168	(0.311)	11	(0.020)	0.399	0.824	0.176	0.218
		C	178	111	(0.624)	61	(0.343)	6	(0.034)		0.795	0.205	

			n	TT		TC		CC			T	C	
rs11610206 *FAM113B*	No E4	AD	404	329	(0.814)	70	(0.173)	5	(0.012)	0.635	0.901	0.099	0.656
		C	729	598	(0.820)	126	(0.173)	5	(0.007)		0.907	0.093	
	E4	AD	553	464	(0.839)	82	(0.148)	7	(0.013)	0.752	0.913	0.087	0.733
		C	173	147	(0.850)	24	(0.139)	2	(0.012)		0.919	0.081	
			n	GG		GA		AA			G	A	
rs11159647 *LOC100287804*	No E4	AD	401	111	(0.277)	204	(0.509)	86	(0.214)	0.842	0.531	0.469	0.755
		C	723	211	(0.292)	356	(0.492)	156	(0.216)		0.538	0.462	
	E4	AD	541	159	(0.294)	277	(0.512)	105	(0.194)	0.333	0.550	0.450	0.457
		C	182	54	(0.297)	84	(0.462)	44	(0.242)		0.528	0.473	
			n	AA		AC		CC			A	C	
rs2049161 *GAPDHL11*	No E4	AD	336	229	(0.682)	92	(0.274)	15	(0.045)	0.964	0.819	0.182	0.779
		C	663	457	(0.689)	178	(0.268)	28	(0.042)		0.824	0.177	
	E4	AD	441	293	(0.664)	129	(0.293)	19	(0.043)	0.054	0.811	0.189	0.172
		C	176	131	(0.744)	35	(0.199)	10	(0.057)		0.844	0.156	
			n	GG		GA		AA			G	A	
rs3746319 *ZNF224*	No E4	AD	414	276	(0.667)	126	(0.304)	12	(0.029)	0.900	0.819	0.181	0.941
		C	767	516	(0.673)	226	(0.295)	25	(0.033)		0.820	0.180	
	E4	AD	566	379	(0.670)	166	(0.293)	21	(0.037)	0.299	0.816	0.184	0.200
		C	192	122	(0.635)	58	(0.302)	12	(0.063)		0.787	0.214	
			n	AA		AG		GG			A	G	
rs3826656 *CD33*	No E4	AD	348	213	(0.612)	112	(0.322)	23	(0.066)	0.452	0.773	0.227	0.865
		C	699	414	(0.592)	248	(0.355)	37	(0.053)		0.770	0.230	
	E4	AD	463	274	(0.592)	161	(0.348)	28	(0.060)	0.166	0.766	0.234	0.073
		C	175	89	(0.509)	73	(0.417)	13	(0.074)		0.717	0.283	
			n	TT		TC		CC			T	C	
rs5984894 *PCDH11X*	No E4	AD	279	79	(0.283)	138	(0.495)	62	(0.222)	0.754	0.529	0.471	0.594
		C	465	120	(0.258)	239	(0.514)	106	(0.228)		0.516	0.484	
	E4	AD	378	108	(0.286)	169	(0.447)	101	(0.267)	0.071	0.503	0.497	0.803
		C	116	26	(0.224)	66	(0.569)	24	(0.207)		0.511	0.489	

3.9 STRATIFICATION BY SEX

Sex stratified data for 15 SNPs are presented in Table 10. One SNP, rs1049951/*PAOX*, showed a statistically significant difference between cases and controls in females (p=0.013).

Table 10. Distribution of 15 SNPs, based on sex

SNP/Gene			*n*	Genotype count (frequency)						*p*
				CC		CT		TT		
rs179943 *ATXN1*	males	AD	288	254	(0.882)	32	(0.111)	2	(0.007)	0.171
		C	378	316	(0.836)	60	(0.159)	2	(0.005)	
	females	AD	619	530	(0.856)	83	(0.134)	6	(0.010)	0.628
		C	536	450	(0.840)	82	(0.153)	4	(0.007)	
				CC		CT		TT		
rs7893928 *CXCL12*	males	AD	323	239	(0.740)	77	(0.238)	7	(0.022)	0.867
		C	385	278	(0.722)	98	(0.255)	9	(0.023)	
	females	AD	666	477	(0.716)	173	(0.260)	16	(0.024)	0.297
		C	587	429	(0.731)	137	(0.233)	21	(0.036)	
				TT		TC		CC		
rs16934131 *KCNMA1*	males	AD	324	174	(0.537)	126	(0.389)	24	(0.074)	0.888
		C	387	214	(0.553)	147	(0.380)	26	(0.067)	
	females	AD	669	389	(0.581)	237	(0.354)	43	(0.064)	0.262
		C	589	318	(0.540)	235	(0.399)	36	(0.061)	
				AA		AC		CC		
rs7894737 *SORCS*	males	AD	322	111	(0.345)	155	(0.481)	56	(0.174)	0.215
		C	389	159	(0.409)	168	(0.432)	62	(0.159)	
	females	AD	657	242	(0.368)	314	(0.478)	101	(0.154)	0.926
		C	587	210	(0.358)	286	(0.487)	91	(0.155)	
				TT		TC		CC		
rs11244841 *ADAM12*	males	AD	323	167	(0.517)	137	(0.424)	19	(0.059)	0.618
		C	379	209	(0.551)	147	(0.388)	23	(0.061)	
	females	AD	657	366	(0.557)	245	(0.373)	46	(0.070)	0.798
		C	586	334	(0.570)	216	(0.369)	36	(0.061)	
				AA		AG		GG		
rs1049951 *PAOX*	males	AD	319	289	(0.906)	27	(0.085)	3	(0.009)	0.632
		C	391	349	(0.893)	40	(0.102)	2	(0.005)	
	females	AD	672	582	(0.866)	85	(0.126)	5	(0.007)	0.013
		C	580	526	(0.907)	54	(0.093)	0	(0.000)	
				GG		GA		AA		
rs11599284 *VENTX2*	males	AD	321	251	(0.782)	64	(0.199)	6	(0.019)	0.957
		C	391	308	(0.788)	75	(0.192)	8	(0.020)	
	females	AD	671	504	(0.751)	155	(0.231)	12	(0.018)	0.316
		C	582	458	(0.787)	116	(0.199)	8	(0.014)	
				CC		CT		TT		
rs12781609 *C10orf93*	males	AD	319	131	(0.411)	146	(0.458)	42	(0.132)	0.476
		C	394	167	(0.424)	165	(0.419)	62	(0.157)	
	females	AD	663	257	(0.388)	303	(0.457)	103	(0.155)	0.997
		C	580	224	(0.386)	265	(0.457)	91	(0.157)	
				GG		GT		TT		
rs2373115 *GAB2*	males	AD	303	198	(0.653)	98	(0.323)	7	(0.023)	0.617
		C	363	237	(0.653)	113	(0.311)	13	(0.036)	
	females	AD	620	405	(0.653)	197	(0.318)	18	(0.029)	0.892

		C	552	364	(0.659)	170	(0.308)	18	(0.033)	
				TT		TC		CC		
rs11610206 *FAM113B*	males	AD	316	255	(0.807)	54	(0.171)	7	(0.022)	0.151
		C	360	301	(0.836)	57	(0.158)	2	(0.006)	
	females	AD	641	538	(0.839)	98	(0.153)	5	(0.008)	0.713
		C	549	451	(0.821)	93	(0.169)	5	(0.009)	
				GG		GA		AA		
rs11159647 *LOC100287804*	males	AD	303	77	(0.254)	163	(0.538)	63	(0.208)	0.056
		C	368	119	(0.323)	165	(0.448)	84	(0.228)	
	females	AD	639	193	(0.302)	318	(0.498)	128	(0.200)	0.546
		C	545	149	(0.273)	280	(0.514)	116	(0.213)	
				AA		AC		CC		
rs2049161 *GAPDHL11*	males	AD	239	159	(0.665)	70	(0.293)	10	(0.042)	0.261
		C	350	247	(0.706)	83	(0.237)	20	(0.057)	
	females	AD	538	363	(0.675)	151	(0.281)	24	(0.045)	0.666
		C	490	342	(0.698)	130	(0.265)	18	(0.037)	
				GG		GA		AA		
rs3746319 *ZNF224*	males	AD	318	204	(0.642)	101	(0.318)	13	(0.041)	0.704
		C	382	251	(0.657)	112	(0.293)	19	(0.050)	
	females	AD	662	451	(0.681)	191	(0.289)	20	(0.030)	0.878
		C	584	390	(0.668)	176	(0.301)	18	(0.031)	
				AA		AG		GG		
rs3826656 *CD33*	males	AD	256	144	(0.563)	92	(0.359)	20	(0.078)	0.120
		C	370	223	(0.603)	132	(0.357)	15	(0.041)	
	females	AD	555	343	(0.618)	181	(0.326)	31	(0.056)	0.121
		C	505	281	(0.556)	189	(0.374)	35	(0.069)	
				TT		TC		CC		
rs5984894 *PCDH11X*	males*	AD	320	159	(0.497)	0	(0.000)	161	(0.503)	0.522
		C	378	197	(0.521)	0	(0.000)	181	(0.479)	
	females	AD	657	187	(0.285)	307	(0.467)	163	(0.248)	0.134
		C	588	147	(0.250)	308	(0.524)	133	(0.226)	

*male genotype counts represent hemizygous individuals; *p*-value is that of allele based test

3.10 LINEAR REGRESSION OF SNPS FOR AGE-AT-ONSET (AAO)

Linear regression analysis was performed on APOE-adjusted AAO for each SNP. Results are show in Table 11. Linear regression analysis of APOE-adjusted AAO showed significant associations with rs16934131 in the *KCNMA1* gene on 10q22.3 (p=0.045) and rs3746319 in

ZNF224 gene on 19q13.3 (*p*=0.002). *APOE*-adjusted mean average AAO for *ZNF224*/rs3746319 genotypes GG, GA, and AA were 73.26 ± (SD) 5.9 (n=645), 72.08 ± 6.5 (n=288), and 71.15 ± 6.1 (n=31) years, respectively. When using a recessive model, *KCNMA1*/rs16934131 had a p-value of 0.007. *APOE*-adjusted mean average AAO for *KCNMA1*/rs16934131 genotypes TT + TC and CC were 73.0 ± (SD) 6.1 (n=912) and 70.9 ± 5.7 (n=65) years, respectively.

Table 11. Linear regression of SNPs for AAO, adjusted for *APOE*

SNP	**Gene**	***n***	***p*-value**
rs179943	*ATXN1*	897	0.860
rs7893928	*CXCL12*	974	0.502
rs16934131	*KCNMA1*	977	**0.045**
rs7894737	*SORCS*	963	0.457
rs11244841	*ADAM12*	965	0.509
rs1049951	*PAOX*	980	0.539
rs11599284	*VENTX2*	982	0.261
rs12781609	*C10orf93*	971	0.931
rs2373115	*GAB2*	908	0.645
rs11610206	*FAM113B*	941	0.912
rs11159647	*LOC100287804*	926	0.231
rs2049161	*GAPDHL11*	766	0.631
rs3746319	*ZNF224*	964	**0.002**
rs3826656	*CD33*	800	0.170
rs5984894	*PCDH11X*	961	0.833

3.11 KAPLAN-MEIER ANALYSIS OF AGE AT ONSET

Kaplan-Meier analysis was performed on AAO for two SNPs, *KCNMA1*/rs16934131 and *ZNF224*/rs3746319, on all subjects and stratified by *APOE*4* carrier status. All p-values for Kaplan-Meier analysis can be found in Table 12. *KCNMA1*/rs16934131 was found to be statistically significant (*p*=0.002, *n*=993), particularly among non-*APOE*4* carriers (*p*=0.001,

n=423). rs3746319 was found to be statistically significant among *APOE*4* carriers (*p*=0.036, *n*=566). Kaplan-Meier plots for both SNPs are shown in Figure 8.

Table 12. Kaplan-Meier analysis p-values of rs16934131 and rs3746319

SNP		***n***	***p***
rs16934131 *KCNMA1*	all cases	993	**0.002**
	no E4	423	**0.001**
	E4	570	0.138
rs3746319 *ZNF224*	all cases	980	0.165
	no E4	414	0.923
	E4	566	**0.036**

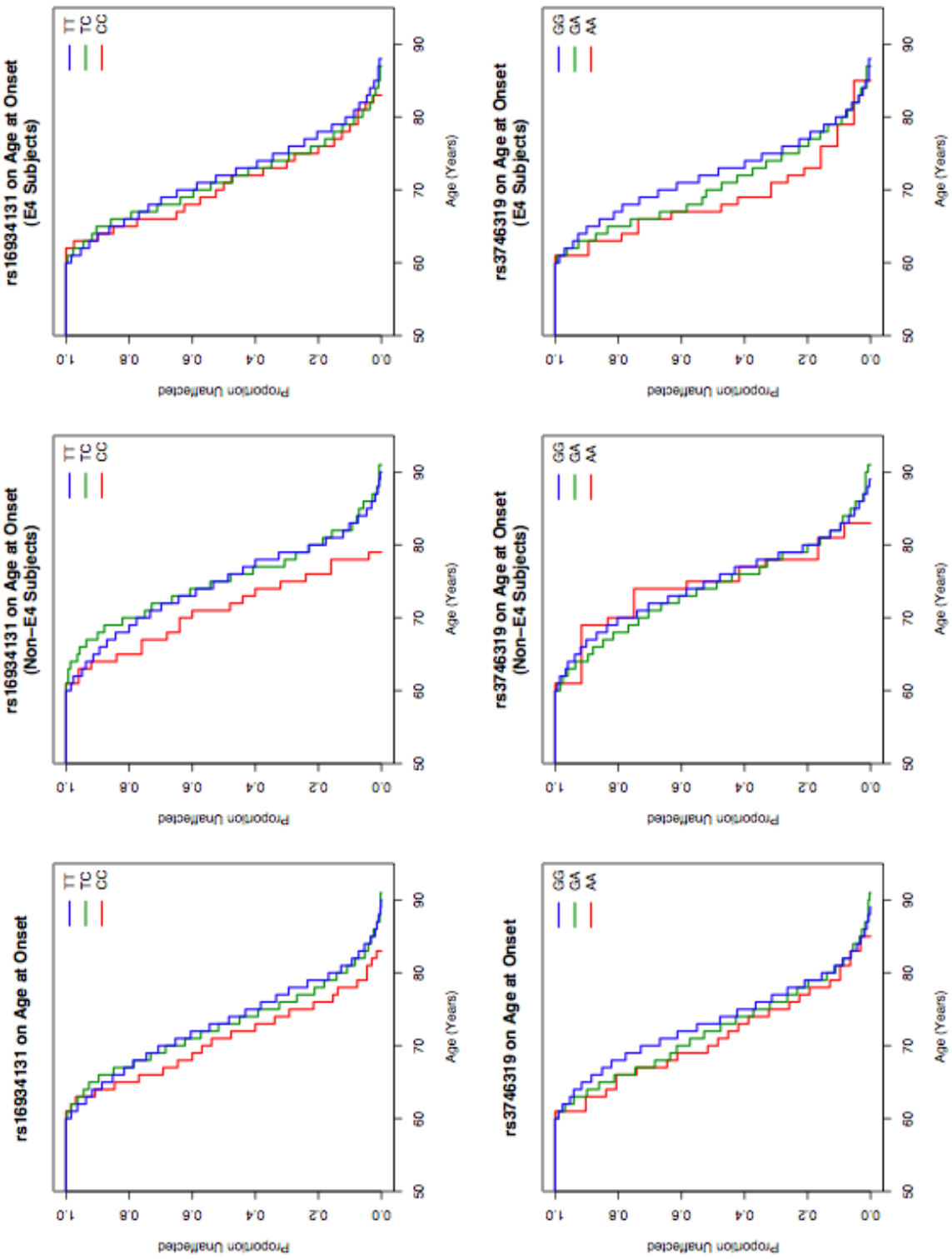

Figure 8. Kaplan-Meier analysis of rs16934131 and rs3746319

3.12 REGRESSION ANALYSIS MODELING FOR INTERACTIONS OF SNPS IN AGE AT ONSET

Regression analysis was performed on AAO for two SNPs, *KCNMA1*/rs16934131 and *ZNF224*/rs3746319. Both SNPs were modeled as additive dominant or recessive, with or without interactions between *APOE*4*. The regression coefficients and p-values for the model with the best fit can be found in Table 13.

Table 13. Regression coefficients and p-values for best-fit regression model

Covariate	Coefficient	p-value
*APOE*4*	-2.618	8.84×10^{-8}
rs16934131 (recessive)	-3.411	0.008
rs3746319 (additive)	-0.285	0.613
*APOE*4* x rs16934131	2.422	0.138
*APOE*4* x rs3746319	-1.294	0.079

3.13 LINEAR REGRESSION OF SNPS FOR DISEASE DURATION

Linear regression analysis was performed on AAO- and sex-adjusted disease duration for each SNP. Results are shown in Table 14. Disease duration data was available on up to 86 AD cases. AAO and gender-adjusted disease duration showed significant association with *KCNMA1*/rs16934131 (p=0.0002) when using a dominant model. This SNP remained significant when taking the number of tests performed into account. Average disease duration for *KCNMA1*/rs16934131 genotypes TT and TC/CC, were 7.65 ± (SD) 3.6 (n=47) and 10.97 ± 3.98 (n=35), respectively.

Table 14. Linear regression of SNPs for disease duration, adjusted for AAO and sex

SNP	Gene	*n*	*p*-value
rs179943	*ATXN1*	80	0.361
rs7893928	*CXCL12*	82	0.864
rs16934131	*KCNMA1*	82	**0.001**
rs7894737	*SORCS*	79	0.637
rs11244841	*ADAM12*	83	0.074
rs1049951	*PAOX*	84	0.804
rs11599284	*VENTX2*	86	0.769
rs12781609	*C10orf93*	84	0.656
rs2373115	*GAB2*	73	0.892
rs11610206	*FAM113B*	78	0.913
rs11159647	*LOC100287804*	78	0.916
rs2049161	*GAPDHL11*	65	0.198
rs3746319	*ZNF224*	82	0.632
rs3826656	*CD33*	67	0.203
rs5984894	*PCDH11X*	80	0.709

3.14 LINEAR REGRESSION OF SNPS FOR BASELINE MINI MENTAL STATE EXAM (MMSE) SCORE

Linear regression analysis was performed on age- and sex-adjusted baseline MMSE score for each SNP. This analysis was performed on cases and controls separately. Results are shown in Table 14. Three SNPs, *KCNMA1*/rs16934131, *C10orf93*/rs12781609, and *PCDH11X/*rs5984894, showed statistically significant association with MMSE score in controls. The most statistically significant SNP was rs16934131 (p=0.002, n=896). Average MMSE score for *KCNMA1*/rs16934131 genotypes TT, TC, and CC were 28.52 ± (SD) 1.36 (n=484), 28.32 ± 1.42 (n=355), and 27.96 ± 1.58 (n=57), respectively. Average MMSE score for *C10orf93/*rs12781609 genotypes CC, CT, and TT were 28.51 ± 1.40 (n=356), 28.36 ± 1.44 (n=393), 28.16 ± 1.44

(n=134), respectively. Average MMSE score for *PCDH11X*/rs5984894 genotypes TT, TC, and CC were 28.50 ± 1.40 (n=308), 28.50 ± 1.40 (n=296), 28.16 ± 1.45 (n=287), respectively.

Table 15. Linear regression of MMSE, adjusted for age and sex

SNP	risk allele		n	P
rs179943 *ATXN1*	T	AD	887	0.111
		C	824	0.628
rs7893928 *CXCL12*	T	AD	968	0.183
		C	889	0.345
rs16934131 *KCNMA1*	C	AD	971	0.502
		C	896	**0.002**
rs7894737 *SORCS*	C	AD	961	0.321
		C	893	0.987
rs11244841 *ADAM12*	C	AD	958	0.425
		C	886	0.487
rs1049951 *PAOX*	G	AD	970	0.990
		C	882	0.283
rs11599284 *VENTX2*	A	AD	971	0.362
		C	882	0.698
rs12781609 *C10orf93*	T	AD	962	0.170
		C	883	**0.012**
rs2373115 *GAB2*	T	AD	905	0.346
		C	834	0.758
rs11610206 *FAM113B*	C	AD	935	0.075
		C	828	0.937
rs11159647 *LOC100287804*	A	AD	922	0.190
		C	833	0.992
rs2049161 *GAPDHL11*	C	AD	760	0.345
		C	752	0.399
rs3746319 *ZNF224*	A	AD	959	0.891
		C	884	0.247
rs3826656 *C33*	G	AD	794	0.692
		C	788	0.070
rs5984894 PCDH11X	C	AD	956	0.456
		C	883	**0.009**

4.0 DISCUSSION

LOAD is a multifactorial condition in which several environmental and genetic risk factors contribute to the etiology of the disease. There are several hypothesized environmental risk factors for LOAD, including education level, gender, tobacco and alcohol use, head trauma, cerebrovascular disease, and cholesterol levels. To date, the only known significant genetic risk factor for LOAD is *APOE*, of which there are three common alleles: *APOE*2*, *APOE*3*, and *APOE*4*. The *APOE*2* allele seems to be protective against AD, while the *APOE*4* allele significantly increases an individual's risk. Although *APOE* is the only confirmed genetic risk factor for LOAD, not all cases carry an *APOE* risk allele, and *APOE* does not explain the entire heritability of LOAD (Kamboh 2004). Therefore, researchers continue to search for additional contributing genetic factors in LOAD.

In order to identify the remaining genes for LOAD, efforts have been focused on a contemporary approach of GWAS, which is a hypothesis free approach and has the potential to identify all known and unknown genes. Recently, six GWAS have implicated additional genes for LOAD (Beecham et al. 2009, Bertram et al. 2008, Carrasquillo et al. 2009, Harold et al. 2009, Lambert et al. 2009, Reiman et al. 2007). Since replication in independent studies is a gold standard before declaring an association being genuine, we have sought to replicate these associations in a large case-controls cohort. We have already replicated findings from two GWAS (Harold et al. 2009, Lambert et al. 2009) in our sample (Kamboh et al. 2010). The focus

of this study was to replicate 12 significant SNPs implicated in four GWAS (Beecham et al. 2009, Bertram et al. 2008, Carrasquillo et al. 2009, Harold et al. 2009, Lambert et al. 2009, Reiman et al. 2007). We also screened 10 additional SNPs present in positional and biological candidate genes located on chromosome 10. In addition to disease risk, we also investigated the associations of these SNPs with AAO, disease duration, and MMSE score.

We did not observe a statistically significant associations between the SNPs and the risk of AD in our primary analysis. Our data suggest that the association in the previously identified SNPs, if it exists, is not statistically significant in our study population. We were unable to replicate the reported GWAS signals in our large case-control samples. Additionally, the candidate SNPs from chromosome 10 also showed no association with AD risk. Furthermore, stratification by *APOE*4* carrier status and stratification by sex failed to reveal any statistically significant associations between SNPs and AD risk.

There are several possible explanations for failing to replicate the reported GWAS associations. One possible explanation is that there are differences between the populations used in the initial studies and the population used in our study. The findings of the various GWAS may in fact be reflecting environmental covariables present in the initial studies' populations that were not present in our study population. Furthermore, heterogeneity in the sample populations may explain the discrepancies between our data and the results found by the various GWAS. The subjects used in this study were Caucasian, while Bertram et al.'s study population consisted of individuals who were of self-reported European ancestry and some individuals in their follow-up analyses were African American. The Carrasquillo et al. study included subjects that were of Caucasian and African American descent. The Beecham et al. study population consisted of individual's of European descent. The term "Caucasian" refers to a large, heterogenous group of

individuals who may be descendants from various regions and countries. Therefore, while the populations used in our study and some of the original studies were primarily made up of Caucasian/European individuals, there could be substantial genetic heterogeneity between the subjects.

Another possible explanation for the associations found by the original studies that were not replicated in our study could be the Winner's Curse. The Winner's Curse refers to a theory that in studies involving many tests, particularly in GWAS, the necessary high level of significance makes it likely that the first study to report a significant finding will also have an effect size much larger than is likely to be seen in replication studies. These "false positives", will therefore not be replicated in subsequent studies, as these studies findings will tend to regress towards the mean (Nakaoka and Inoue 2009).

While none of the SNPs examined in this study were found to be associated with AD risk, we observed some associations with AAO, disease duration, and MMSE score.

Linear regression analysis of *APOE*-adjusted AAO showed significant associations with rs16934131 in the *KCNMA1* gene on 10q22.3 (p=0.045) and rs3746319 in *ZNF224* gene on 19q13.3 (p=0.002). There was a gene-dosage effect of the *ZNF224*/rs3746319 SNP on AAO with the GG genotype with the highest value (73.26 years), the AA genotype with the lowest value (71.15 years) and the GA genotype with the intermediate value (72.08 years). Kaplan-Meier analysis showed that the *ZNF224/*rs3746319 effect on AAO was mainly confined to *APOE*4* carriers (p=0.036, n=566) as reflected by Figure 8 that shows obvious differences between the curve for the AA genotype as compared to those for the GA and GG. These data suggest that the A allele of *ZNF224*/rs3746319 is associated with earlier AAO.

For the *KCNMA1*/rs16934131 SNP, the CC genotype was associated with early AAO (70.88 years) than the TT and TC genotypes (73.00 years). Kaplan-Meier analysis of AAO showed that rs16934131 was found to be statistically significant (p=0.002, n=993), particularly among non-*APOE*4* carriers (p=0.001, n=423). Kaplan-Meier plots seen in Figure 8 show obvious differences between the curve of the CC genotype as compared to those of the TT and TC genotypes, particularly in non-*APOE*4* carriers. The fact that we are seeing a stronger effect of rs16934131 in non-*APOE*4* carriers seems biologically plausible considering that *APOE* is known to have such a strong effect, it may mask other effects. These data suggest that the C allele of *KCNMA1*/rs16934131 is associated with earlier AAO.

The differences between effects of these two SNPs when stratified by *APOE*4* status is suggestive of interactions between these two genes and *APOE*. However, regression analysis did not reveal any statistically significant evidence for interactions between *KCNMA1*/rs16934131, *ZNF224*/rs3746319, and/or *APOE*. This may be partially explained by the fact that when running as many tests as is necessary in regression analysis, the power drops significantly.

A study published by Li et al. (2002) identified potential linkage peaks for AD AAO on chromosome 10. One of these peaks was at 10q, which is also where *KCNMA1*/rs16934131 is located. Therefore, this gene may represent an actual risk factor or may be in linkage disequilibrium with a risk factor for AAO.

The AAO of an individual with AD is dependent upon many factors. One of these factors is the method of diagnosis. As some individuals may have varying degrees of lucidity during their disease, a test of cognitive impairment may not detect all individuals at the exact onset of their disease. The diagnosis of AD is also partially depending on report of cognitive decline by an individual's peers and/or family members, which are subject to biases as well. If an individual

with developing AD either has no family members that he interacts with, or these family members fail to notice the signs, they may be diagnosed at a later age (Cruz et al. 2004). It is important to keep these factors in mind when interpreting these results.

In addition to the association of the *KCNMA1*/rs16934131 SNP with AAO, it also revealed an association with disease duration (p=0.0002). The C allele (TC/CC genotypes) was associated with longer disease duration (10.97 years) than TT homozygotes (7.65 years). Since there were only two individuals with the CC genotype, the main contribution for longer disease duration was provided by the TC heterozygosity. We cannot comment on the effect of being CC at the *KCNMA1*/rs16934131 locus because disease duration information was only available for two individuals with this genotype.

It is important to note that this analysis is missing vital information regarding disease duration. One piece of information that could aid in the interpretation of these results is cause of death of subjects. According to the Alzheimer's Association, the average disease duration for individuals diagnosed with AD after age 60 years is 4-6 years (Alzheimer's Association 2009). Although AD can lead to death, AD subjects may die of various causes. However, it is necessary to distinguish between death *with* AD and death *from* AD. Severe dementia can cause complications such as difficulty with mobility, swallowing disorders, and malnutrition. These complications significantly increase the risk for pneumonia, which is the most frequently identified cause of death in individuals with AD (Alzheimer's Association 2009). Additionally, many elderly individuals die of causes that precipitate from one single event. For example, hip fracture is one of the most common injuries in the elderly. Hip injury often results from loss of mobility and balance, which can be features associated with AD. It is estimated that elderly individuals have a 5- to 8-fold increase in mortality following hip fracture (Haentjens et al.

2010). Therefore it is important to consider that there are causes of death, such as hip fracture, that may or may not be linked to AD.

Another factor that may impact an individual's disease duration is their quality of care and/or socioeconomic status, which would also help in the interpretation of these results. Considering AD is a debilitating disease, an individual's quality of care may greatly impact the length of time that they live with AD. Without adequate care, individuals with AD may die of various other conditions that result from the inability to care for oneself. Socioeconomic status can also contribute, because an individual with AD who has the financial means to pay for an assisted living center and/or nursing home will likely live longer than an impoverished individual with AD who must care for themselves.

Considering these factors and various others that may contribute to disease duration, it is possible that these data represent a confounding factor. Perhaps this SNP actually segregates with financial success, which leads to an individual having better access to care later in life, which leads to longer disease duration. Or possibly this SNP actually has an impact on an individual's balance, and therefore individuals with a certain genotype have a greater propensity towards falling and breaking their hip which may ultimately lead to their death. Without additional information on cause of death and other factors it is impossible to hypothesize how this SNP may impact disease duration.

The two SNPs that were found in our study to be associated with AAO and/or disease duration, *KCNMA1*/rs16934131 and *ZNF224*/rs3746319, were originally found to be associated with AD risk by Beecham et al. (2009). While our study did not find associations between these SNPs and AD risk, the associations seen in our study may be consistent with Beecham et al.'s findings. If an individual has a genetic factor that predisposes them to have a longer disease

duration, they are living longer with AD, and therefore there may be a greater chance that this individual would be chosen as a case rather than a control. Similarly, if an individual has a genetic factor that causes an earlier AAO, they would be affected with AD at an earlier age which could also increase the likelihood that they would be classified as an AD case. Therefore, depending on how old an individual is when they are enrolled in a study, what was seen in our study as a difference in disease duration or AAO, may be reflected in another study as a difference in case/control status.

The *KCNMA1*/rs16934131 also revealed association with MMSE score in controls, in addition to the *C10orf93*/rs12781609, and *PCDH11X*/rs5984894 SNPs. The *KCNMA1*/rs16934131 SNP was the most statistically significant (p=0.002, n=896).

Since MMSE score reflects cognitive impairment, it is possible that these controls actually do have some degree of cognitive impairment but not enough to be diagnosed with AD. Depending on the method of diagnosis, this difference in MMSE score could be reflected as difference in case-control status. It is possible that individuals with a particular genotype at this SNP have a greater degree of cognitive impairment and would be diagnosed as having AD when more stringent criteria for diagnosis is used. Therefore, potentially the same subjects that were classified as cases in other studies because of their cognitive impairment were classified as controls in our study, yet their cognitive impairment was reflected as a lower MMSE score.

It is interesting to note that *KCNMA1*/rs16934131 was found to be associated with AAO, disease duration, and MMSE score in this study. A summary of the effects seen associated with *KCNMA1*/rs16934131 are summarized in Table 16. It is possible that this SNP does impact all three measures independently, or perhaps that it has one common underlying effect that is reflected in all three measures. It seems plausible that there could be a genetic factor that would

cause greater cognitive impairment, which would be represented as a lower MMSE score, particularly in individuals who have not had a diagnosis of AD (controls). This same effect of greater cognitive impairment may cause individuals to be diagnosed with AD at an earlier age, thus causing an overall earlier AAO. Additionally, if this genetic factor does not cause the AD to progress any more rapidly and these individuals would still die at a particular age, this earlier AAO would lead to a longer time between diagnosis and death, thus a longer disease duration.

Table 16. Summary of effects associated with rs16934131

rs16934131 Genotype	Means		
	AAO (years)	Disease Duration (years)	MMSE score
TT	73.06	7.65	28.52
TC	72.93	10.97	28.32
CC	70.88		27.96
p-value	0.045	0.0002	0.002

The *KCNMA1* gene has been suggested to be associated with innate immunity, cochlear hearing, vascular dysfunction, epilepsy, and mental retardation. Knockout mice have been engineered by targeted disruption of the *KCNMA1* gene which showed obvious ataxia, moderate vascular dysfunction, and independent neurologic deficits, including cerebellar dysfunction in the form of abnormal eye-blink reflexes, abnormal motion, and deficiency in motor coordination (Sausbier et al. 2004). Considering the evidence that this gene functions in the brain and could potentially impact mental capacity, it seems biologically plausible that alterations in this gene could have an effect on an individual's degree of cognitive impairment. Additionally, in light of evidence that this gene may play a role in motor coordination, it is also possible that this SNP's effect on disease duration actually represents a confounding factor, such as an effect on balance, as previously discussed.

It is important to consider the implications of these findings. Although the contribution of these implicated SNPs to AD risk and aspects of AD may be small, these findings may help us more fully understand the biology behind AD, and may lead to more effective treatments in the future. Considering the financial impact of AD, if preventative measures could be taken in order to delay the onset of the disease, the financial burden on the health care system could be significantly lessened. When considering the large financial and social burden of AD, even small reductions in risks may have major public health significance.

In summary, we did not observe statistically significant associations between the SNPs and the risk of AD, therefore we were unable to replicate the reported GWAS signals in our large case-control samples. Additionally, the candidate SNPs from chromosome 10 also showed no association with AD risk. Stratification by *APOE*4* carrier status and stratification by sex failed to reveal any statistically significant associations between SNPs and AD risk. However, additional analysis to examine the potential associations between the examined SNPs and AAO, disease duration, and baseline MMSE score revealed several SNPs with statistically significant associations. One SNP, *KCNMA1*/rs16934131, was found to be associated with AAO, disease duration, and MMSE score.

We must study these genes in much greater depth in order to examine these SNPs and all others with potential biological or positional significance. Future research may elucidate risk alleles for late onset AD.

BIBLIOGRAPHY

Alzheimer's Association. 2009. Alzheimer's Disease Facts & Figures.

Alzheimer Disease & Frontotemporal Dementia Mutation Database. 2010. http://www.molgen.ua.ac.be/Admutations. Accessed 28 February 2010.

Bassett SS, Avramopoulos D, Fallin D. 2002. Evidence for parent of origin effect in late-onset Alzheimer disease. Am J Med Genet 114: 679-686.

Beecham GW, Martin ER, Li YJ, Slifer MA, Gilbert JR, Haines JL, Pericak-Vance MA. 2009. Genome-wide association study implicates a chromosome 12 risk locus for late-onset Alzheimer disease. Am J Hum Genet 84: 35-43.

Bennett DA, Schneider JA, Wilson RS, Bienias JL, Arnold SE. 2005. Education modifies the association of amyloid but not tangles with cognitive function. Neurology 65: 953-955.

Bentires-Alj M, Gil SG, Chan R, Wang ZC, Wang Y, Imanaka N, Harris LN, Richardson A, Neel BG, Gu H. 2006. A role for the scaffolding adapter GAB2 in breast cancer. Nat Med 12: 114-121.

Bertram L, Lange C, Mullin K, Parkinson M, Hsiao M, Hogan MF, Schjeide BM, Hooli B, Divito J, Ionita I, Jiang H, Laird N, Moscarillo T, Ohlsen KL, Elliott K, Wang X, Hu-Lince D, Ryder M, Murphy A, Wagner SL, Blacker D, Becker KD, Tanzi RE. 2008. Genome-wide association analysis reveals putative Alzheimer's disease susceptibility loci in addition to APOE. Am J Hum Genet 83: 623-632.

Borgaonkar DS, Schmidt LC, Martin SE, Kanzer MD, Edelsohn L, Growdon J, Farrer LA. 1993. Linkage of late-onset Alzheimer's disease with apolipoprotein E type 4 on chromosome 19. Lancet 342: 625.

Carrasquillo MM, Zou F, Pankratz VS, Wilcox SL, Ma L, Walker LP, Younkin SG, Younkin CS, Younkin LH, Bisceglio GD, Ertekin-Taner N, Crook JE, Dickson DW, Petersen RC, Graff-Radford NR. 2009. Genetic variation in PCDH11X is associated with susceptibility to late-onset Alzheimer's disease. Nat Genet 41: 192-198.

Cruz VT, Pais J, Teixeira A, Nunes B. 2004. [The initial symptoms of Alzheimer disease: caregiver perception]. Acta Med Port 17: 435-444.

Cummings CJ, Reinstein E, Sun Y, Antalffy B, Jiang Y, Ciechanover A, Orr HT, Beaudet AL, Zoghbi HY. 1999. Mutation of the E6-AP ubiquitin ligase reduces nuclear inclusion frequency while accelerating polyglutamine-induced pathology in SCA1 mice. Neuron 24: 879-892.

Ertekin-Taner N. 2007. Genetics of Alzheimer's disease: a centennial review. Neurol Clin 25: 611-667, v.

Filley CM, Rollins YD, Anderson CA, Arciniegas DB, Howard KL, Murrell JR, Boyer PJ, Kleinschmidt-DeMasters BK, Ghetti B. 2007. The genetics of very early onset Alzheimer disease. Cogn Behav Neurol 20: 149-156.

Folstein MF, Folstein SE, McHugh PR. 1975. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 12: 189-198.

Galliano MF, Huet C, Frygelius J, Polgren A, Wewer UM, Engvall E. 2000. Binding of ADAM12, a marker of skeletal muscle regeneration, to the muscle-specific actin-binding protein, alpha -actinin-2, is required for myoblast fusion. J Biol Chem 275: 13933-13939.

Gatz M, Reynolds CA, Fratiglioni L, Johansson B, Mortimer JA, Berg S, Fiske A, Pedersen NL. 2006. Role of genes and environments for explaining Alzheimer disease. Arch Gen Psychiatry 63: 168-174.

Haentjens P, Magaziner J, Colon-Emeric CS, Vanderschueren D, Milisen K, Velkeniers B, Boonen S. 2010. Meta-analysis: excess mortality after hip fracture among older women and men. Ann Intern Med 152: 380-390.

Hampe W, Rezgaoui M, Hermans-Borgmeyer I, Schaller HC. 2001. The genes for the human VPS10 domain-containing receptors are large and contain many small exons. Hum Genet 108: 529-536.

Han ZG, Zhang QH, Ye M, Kan LX, Gu BW, He KL, Shi SL, Zhou J, Fu G, Mao M, Chen SJ, Yu L, Chen Z. 1999. Molecular cloning of six novel Kruppel-like zinc finger genes from hematopoietic cells and identification of a novel transregulatory domain KRNB. J Biol Chem 274: 35741-35748.

Harold D, Abraham R, Hollingworth P, Sims R, Gerrish A, Hamshere ML, Pahwa JS, Moskvina V, Dowzell K, Williams A, Jones N, Thomas C, Stretton A, Morgan AR, Lovestone S, Powell J, Proitsi P, Lupton MK, Brayne C, Rubinsztein DC, Gill M, Lawlor B, Lynch A, Morgan K, Brown KS, Passmore PA, Craig D, McGuinness B, Todd S, Holmes C, Mann D, Smith AD, Love S, Kehoe PG, Hardy J, Mead S, Fox N, Rossor M, Collinge J, Maier W, Jessen F, Schurmann B, van den Bussche H, Heuser I, Kornhuber J, Wiltfang J, Dichgans M, Frolich L, Hampel H, Hull M, Rujescu D, Goate AM, Kauwe JS, Cruchaga C, Nowotny P, Morris JC, Mayo K, Sleegers K, Bettens K, Engelborghs S, De Deyn PP, Van Broeckhoven C, Livingston G, Bass NJ, Gurling H, McQuillin A, Gwilliam R, Deloukas P, Al-Chalabi A, Shaw CE, Tsolaki M, Singleton AB, Guerreiro R, Muhleisen TW, Nothen MM, Moebus S, Jockel KH, Klopp N, Wichmann HE, Carrasquillo MM, Pankratz VS, Younkin SG, Holmans PA, O'Donovan M, Owen MJ, Williams J. 2009.

Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. Nat Genet 41: 1088-1093.

Hartmann T, Kuchenbecker J, Grimm MO. 2007. Alzheimer's disease: the lipid connection. J Neurochem 103 Suppl 1: 159-170.

Hermey G, Riedel IB, Hampe W, Schaller HC, Hermans-Borgmeyer I. 1999. Identification and characterization of SorCS, a third member of a novel receptor family. Biochem Biophys Res Commun 266: 347-351.

Kalaria RN. 2003. Comparison between Alzheimer's disease and vascular dementia: implications for treatment. Neurol Res 25: 661-664.

Kamboh MI. 2004. Molecular genetics of late-onset Alzheimer's disease. Ann Hum Genet 68: 381-404.

Kamboh MI, Minster RL, Demirci FY, Ganguli M, DeKosky ST, Lopez OL, Barmada MM. 2010. Association of CLU and PICALM variants with Alzheimer's disease. Neurobiol Aging (In Press).

Kawas C, Resnick S, Morrison A, Brookmeyer R, Corrada M, Zonderman A, Bacal C, Lingle DD, Metter E. 1997. A prospective study of estrogen replacement therapy and the risk of developing Alzheimer's disease: the Baltimore Longitudinal Study of Aging. Neurology 48: 1517-1521.

Kumar V, Kinsella LJ. 2010. Healthy brain aging: effect of head injury, alcohol and environmental toxins. Clin Geriatr Med 26: 29-44.

Lambert JC, Heath S, Even G, Campion D, Sleegers K, Hiltunen M, Combarros O, Zelenika D, Bullido MJ, Tavernier B, Letenneur L, Bettens K, Berr C, Pasquier F, Fievet N, Barberger-Gateau P, Engelborghs S, De Deyn P, Mateo I, Franck A, Helisalmi S, Porcellini E, Hanon O, de Pancorbo MM, Lendon C, Dufouil C, Jaillard C, Leveillard T, Alvarez V, Bosco P, Mancuso M, Panza F, Nacmias B, Bossu P, Piccardi P, Annoni G, Seripa D, Galimberti D, Hannequin D, Licastro F, Soininen H, Ritchie K, Blanche H, Dartigues JF, Tzourio C, Gut I, Van Broeckhoven C, Alperovitch A, Lathrop M, Amouyel P. 2009. Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. Nat Genet 41: 1094-1099.

Lee TH, Lwu S, Kim J, Pelletier J. 2002. Inhibition of Wilms tumor 1 transactivation by bone marrow zinc finger 2, a novel transcriptional repressor. J Biol Chem 277: 44826-44837.

Letenneur L, Larrieu S, Barberger-Gateau P. 2004. Alcohol and tobacco consumption as risk factors of dementia: a review of epidemiological studies. Biomed Pharmacother 58: 95-99.

Letenneur L, Gilleron V, Commenges D, Helmer C, Orgogozo JM, Dartigues JF. 1999. Are sex and educational level independent predictors of dementia and Alzheimer's disease? Incidence data from the PAQUID project. J Neurol Neurosurg Psychiatry 66: 177-183.

Letenneur L, Launer LJ, Andersen K, Dewey ME, Ott A, Copeland JR, Dartigues JF, Kragh-Sorensen P, Baldereschi M, Brayne C, Lobo A, Martinez-Lage JM, Stijnen T, Hofman A. 2000. Education and the risk for Alzheimer's disease: sex makes a difference. EURODEM pooled analyses. EURODEM Incidence Research Group. Am J Epidemiol 151: 1064-1071.

Li YJ, Scott WK, Hedges DJ, Zhang F, Gaskell PC, Nance MA, Watts RL, Hubble JP, Koller WC, Pahwa R, Stern MB, Hiner BC, Jankovic J, Allen FA, Jr., Goetz CG, Mastaglia F, Stajich JM, Gibson RA, Middleton LT, Saunders AM, Scott BL, Small GW, Nicodemus KK, Reed AD, Schmechel DE, Welsh-Bohmer KA, Conneally PM, Roses AD, Gilbert JR, Vance JM, Haines JL, Pericak-Vance MA. 2002. Age at onset in two common neurodegenerative diseases is genetically controlled. Am J Hum Genet 70: 985-993.

Liddell MB, Lovestone S, Owen MJ. 2001. Genetic risk of Alzheimer's disease: advising relatives. Br J Psychiatry 178: 7-11.

Luedi PP, Dietrich FS, Weidman JR, Bosko JM, Jirtle RL, Hartemink AJ. 2007. Computational and experimental identification of novel human imprinted genes. Genome Res 17: 1723-1730.

Mazzola JL, Sirover MA. 2001. Reduction of glyceraldehyde-3-phosphate dehydrogenase activity in Alzheimer's disease and in Huntington's disease fibroblasts. J Neurochem 76: 442-449.

Mejia S, Giraldo M, Pineda D, Ardila A, Lopera F. 2003. Nongenetic factors as modifiers of the age of onset of familial Alzheimer's disease. Int Psychogeriatr 15: 337-349.

Nakaoka H, Inoue I. 2009. Meta-analysis of genetic association studies: methodologies, between-study heterogeneity and winner's curse. J Hum Genet 54: 615-623.

Orgogozo JM, Dartigues JF, Lafont S, Letenneur L, Commenges D, Salamon R, Renaud S, Breteler MB. 1997. Wine consumption and dementia in the elderly: a prospective community study in the Bordeaux area. Rev Neurol (Paris) 153: 185-192.

Panza F, D'Introno A, Colacicco AM, Basile AM, Capurso C, Kehoe PG, Capurso A, Solfrizzi V. 2004. Vascular risk and genetics of sporadic late-onset Alzheimer's disease. J Neural Transm 111: 69-89.

Paradise M, Cooper C, Livingston G. 2009. Systematic review of the effect of education on survival in Alzheimer's disease. Int Psychogeriatr 21: 25-32.

Peiper SC, Ashmun RA, Look AT. 1988. Molecular cloning, expression, and chromosomal localization of a human gene encoding the CD33 myeloid differentiation antigen. Blood 72: 314-321.

Plassman BL, Havlik RJ, Steffens DC, Helms MJ, Newman TN, Drosdick D, Phillips C, Gau BA, Welsh-Bohmer KA, Burke JR, Guralnik JM, Breitner JC. 2000. Documented head

injury in early adulthood and risk of Alzheimer's disease and other dementias. Neurology 55: 1158-1166.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81: 559-575.

Reiman EM, Webster JA, Myers AJ, Hardy J, Dunckley T, Zismann VL, Joshipura KD, Pearson JV, Hu-Lince D, Huentelman MJ, Craig DW, Coon KD, Liang WS, Herbert RH, Beach T, Rohrer KC, Zhao AS, Leung D, Bryden L, Marlowe L, Kaleem M, Mastroeni D, Grover A, Heward CB, Ravid R, Rogers J, Hutton ML, Melquist S, Petersen RC, Alexander GE, Caselli RJ, Kukull W, Papassotiropoulos A, Stephan DA. 2007. GAB2 alleles modify Alzheimer's risk in APOE epsilon4 carriers. Neuron 54: 713-720.

Reiss AB. 2005. Cholesterol and apolipoprotein E in Alzheimer's disease. Am J Alzheimers Dis Other Demen 20: 91-96.

Ronaghi M, Uhlen M, Nyren P. 1998. A sequencing method based on real-time pyrophosphate. Science 281: 363, 365.

Sando SB, Melquist S, Cannon A, Hutton M, Sletvold O, Saltvedt I, White LR, Lydersen S, Aasly J. 2008. Risk-reducing effect of education in Alzheimer's disease. Int J Geriatr Psychiatry 23: 1156-1162.

Sausbier M, Hu H, Arntz C, Feil S, Kamm S, Adelsberger H, Sausbier U, Sailer CA, Feil R, Hofmann F, Korth M, Shipston MJ, Knaus HG, Wolfer DP, Pedroarena CM, Storm JF, Ruth P. 2004. Cerebellar ataxia and Purkinje cell dysfunction caused by Ca2+-activated K+ channel deficiency. Proc Natl Acad Sci U S A 101: 9474-9478.

Servadio A, Koshy B, Armstrong D, Antalffy B, Orr HT, Zoghbi HY. 1995. Expression analysis of the ataxin-1 protein in tissues from normal and spinocerebellar ataxia type 1 individuals. Nat Genet 10: 94-98.

Szczygielski J, Mautes A, Steudel WI, Falkai P, Bayer TA, Wirths O. 2005. Traumatic brain injury: cause or risk of Alzheimer's disease? A review of experimental studies. J Neural Transm 112: 1547-1564.

Team RDC. 2007. R: A Language and Environment for Statistical Computing in Austria Vienna, ed: R Foundation for Statistical Computing.

Vilalta-Franch J, Lopez-Pousa S, Turon-Estrada A, Lozano-Gallego M, Hernandez-Ferrandiz M, Pericot-Nierga I, Garre-Olmo J. 2010. Syndromic Association of Behavioral and Psychological Symptoms of Dementia in Alzheimer Disease and Patient Classification. Am J Geriatr Psychiatry.

Welch HG, Walsh JS, Larson EB. 1992. The cost of institutional care in Alzheimer's disease: nursing home and hospital use in a prospective cohort. J Am Geriatr Soc 40: 221-224.

World Alzheimer Report. 2009. Alzheimer's Disease International.

Yoshida K, Sugano S. 1999. Identification of a novel protocadherin gene (PCDH11) on the human XY homology region in Xq21.3. Genomics 62: 540-543.

Zekanowski C, Religa D, Graff C, Filipek S, Kuznicki J. 2004. Genetic aspects of Alzheimer's disease. Acta Neurobiol Exp (Wars) 64: 19-31.

Zhao C, Yu DH, Shen R, Feng GS. 1999. Gab2, a new pleckstrin homology domain-containing adapter protein, acts to uncouple signaling from ERK kinase to Elk-1. J Biol Chem 274: 19649-19654.