Molecular Genetics of MS4A6A and Alzheimer's Disease

Ryan Harpole
University of Kentucky, ryan.harpole@yahoo.com

Follow this and additional works at: https://uknowledge.uky.edu/honprog

Part of the Genetic Phenomena Commons, Genetic Processes Commons, Genetic Structures Commons, Medical Biochemistry Commons, and the Medical Genetics Commons

Right click to open a feedback form in a new tab to let us know how this document benefits you.

Recommended Citation
Harpole, Ryan, "Molecular Genetics of MS4A6A and Alzheimer's Disease" (2016). Lewis Honors College Capstone Collection. 28.
https://uknowledge.uky.edu/honprog/28

This Article is brought to you for free and open access by the Lewis Honors College at UKnowledge. It has been accepted for inclusion in Lewis Honors College Capstone Collection by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.
I. Project Title

“Molecular Genetics of MS4A6A and Alzheimer's Disease”

II. Pertinent Student Information

Ryan Harpole  
1385 Hickerson Road  
Clinton, KY 42031  
(270) 254-1759  
ryan.harpole@yahoo.com  
Fall, 2017

Signed: ___________________________  Date: ________________

III. Name of Faculty Member who will Supervise the Activity

Dr. Steven Estus  
Sanders Brown Center on Aging  
University of Kentucky

Signed: ___________________________  Date: ________________

IV. Statement of Career Goals

Upon graduation from the University of Kentucky, I plan on attending medical school. Once I have graduated from medical school, I intend on becoming a dermatologist, and opening up a practice in Paducah, KY.
V. Abstract

Increased Alzheimer’s disease (AD) risk has previously been associated with a SNP called rs610932 near the gene MS4A6A. The goal of this experiment was to quantify the expression of two MS4A6A isoforms in the brains of AD and non-AD subjects, particularly as a function of rs610932 genotype. According to an article titled “Alzheimer’s Disease Susceptibility Variants in the MS4A6A Gene are Associated with Altered Levels of MS4A6A Expression in Blood”, MS4A6A has four different isoforms that have been reported to be differentially expressed in the blood of AD subjects compared to non-AD subjects (Petroula et al., 2014). After statistically analyzing the association between two of the MS4A6A variants, Isoform 6 (total MS4A6A transcript) and a variant, Isoform 4, and rs610932 using AD status and microglial expression as independent variables, we found significant expression differences in AD and non-AD subjects in the brain for the Isoform 4. We did not detect significant expression differences for Isoform 6 in AD subjects, nor did we find significant expression differences in either isoform as a function of the rs610932 SNP. These results are in contrast to previous findings that suggested a significant expression difference due to rs610932 in whole blood (Petroula et al., 2014). In summary, further work is necessary to understand how rs610932 modulates MS4A6A to alter AD risk.

VI. Introduction

Nearly 70% of the risk of late onset Alzheimer’s disease (AD) is heritable. AD risk has been linked with genetic variants in different genes. This disease is a form of dementia that affects memory, thinking and behavior. AD tends to start slowly, but then
develops over time. The largest risk factor to acquire AD is age and there is no current cure for the disease. Symptoms that correspond to AD include memory loss, confusion, mood changes, and poor judgment (Malik et al. 2015).

There are two main neuropathologic findings that correlate with AD. The first is when a type of plaque builds between the neurons of the brain (Sakae et al. 2016). These plaques are made of beta-amyloid protein, and they usually arise in the memory area of the brain (Sakae et al. 2016). The second finding is the formation of neurofibrillary tangles (Baker and Gotz 2016). These tangles are composed of a protein called tau that builds up inside of the neurons (Baker and Gotz 2016). Plaques and tangles are believed to block nerve cell communication, eventually leading to cell death (Baker and Gotz 2016).

The root cause of AD is not fully known, but large-scale genetic studies have identified a set of SNPs that are associated with AD risk. These SNPs have suggested that AD is the result of an immune system defect. One gene in particular that has recently emerged as a candidate for contributing to AD is MS4A6A (Deng et al., 2012). This gene is part of the MS4A gene family, which consists of 12 related genes. The gene is found on chromosome 11q12.1. Although there are 12 members of the MS4A gene family, MS4A1, MS4A2, and MS4A3 are the most well known genes (Petroula et al., 2014).

MS4A6A encodes a member of the membrane-spanning 4A (MS4A) gene family, and typically consists of six exons. The proteins produced from this gene are often alternatively spliced, and there are four known isoform transcripts that are translated. A schematic representation of two MS4A6A variants analyzed in this study is shown in Figure 1 (Petroula et al., 2014). MS4A6A is 13,060 base pairs in length and encodes a
total of 248 amino acids in the transcript’s full form. The different isoforms produced from this gene are the result of variations in splicing and, possibly, single nucleotide polymorphisms (SNPs). The SNP of interest is rs610931. This SNP causes a thymine (T) to be changed to a guanine (G), and is located at the end of the MS4A6A gene (Petroula et al., 2014). G is the major allele and T is the minor, protective allele. This SNP has been associated with altered MS4A6A expression in whole blood samples (Petroula et al., 2014).

This particular gene, along with its four variant transcripts, is hypothesized to contribute to AD susceptibility. There is evidence that subjects with Alzheimer’s disease have an increased expression of the MS4A6A transcript in the blood (Petroula et al., 2014). The goal of this experiment was to link two of the MS4A6A variants’ gene expression in the brain, Isoform 6 and Isoform 4 (V1 and V4 in Figure 1), with Alzheimer’s disease, particularly as a result of rs610932.

**VII. Methods**

The conductance of all experiments were in compliance with specific, applicable Chemical Hygiene, Radiation Safety, Institutional Biological Safety, and Institutional Review Board protocols.
Subjects:

The University of Kentucky Alzheimer’s Disease Center Neuropathology Core provided all human brain specimens. The specimens came from the anterior cingulate. Figure 2 shows a diagram of where the anterior cingulate is located in the brain (Thompson 2013). Diagnosing the AD and non-AD subjects was done at a conference of the AD Center Neuropathology and Clinical Cores. Diagnosis was based on evaluation of cognitive status and neuropathology. Cognitive status was determined using tests such as Clinical Dementia Rating and Mental State Examination (MMSE) scores (Grear et al., 2009). Neuropathology was determined based on tests such as Braak stages and NIA-Reagan Institute (NIA-RI) neuropathological classification (Grear et al., 2009). NIA-RI includes counts of both neurofibrillary tangles and neuritic senile plaques, and provides staging for AD diagnosis. Individuals that were diagnosed non-AD had an age of death of 82 ± 9 years (mean ± SD, n = 28). Those diagnosed as AD had an age of death of 82 ± 6 years (n=29). In non-AD patients, the average post mortem interval (PMI) was 2.8 ± 0.8 hours (mean ± SD, n = 28), while AD individuals had a PMI of 3.4 ± 0.6 hours (n = 29). Mini-Mental State Examination (MMSE) scores for non-AD individuals were 28.4 ±1.6 (n = 28). The temporal lobe samples were prepared in order to compare MS4A6A expression between white and gray matter. White and gray matter was removed from the temporal lobe, and the quality of these specimens was confirmed by comparing the ratio of mRNAs associated with different cell types (Grear et al., 2009).
Genotyping:

The first method used in this study was to genotype the genomic DNA samples with a SNP kit that corresponds to the rs610932 SNP found in the MS4A6A gene. The TaqMan® SNP Genotyping Assays were ordered from Applied Biosystems (Foster City, California, USA). They were centrifuged and diluted with 564 μL of water. A master mix was then made using 700 μL of TaqMan® Genotyping Master Mix, 70 μL of the diluted SNP Kit, and 560 μL of water. The TaqMan® Genotyping MM was also ordered from Applied Biosystems. Nineteen μL of the master mix was pipetted into target wells, followed by 1 μL of DNA from the samples. Negative controls that had no DNA were also used. The pipettes used in this experiment (Pipetmen, Gilson Inc., Middleton, Wisconsin, USA) were used with ZAP Premier sterile aerosol pipet tips (Labcon Inc., Petaluma, CA, USA) to eliminate aerosol carryover between samples. Once the wells had all 20 μL of liquid, the 96-well plate was placed in an MJ Research Peltier Thermal Cycler (PTC-200, MJ Research, Bio-Rad Inc., Hercules, CA, USA). The plate was incubated at 95° C for 10 minutes, 95° C for 15 seconds, and 60° C for 1 minute. Steps 2 and 3 were repeated for 39 more minutes. Once the 96-well plate was incubated, the results were analyzed through MJ Opticon Monitor Analysis Software (version 3.1; Bio-Rad Inc.). The graph showing the genotyping results is shown in Figure 3.

Test PCR Primers:

Primers for MS4A6A were ordered (Invitrogen Inc., Carlsbad, CA, USA) for exons 1, 2, 4, 5, and 6. The primers were first dissolved with water to generate a stock concentration of 50 μM. Once done, a master mix was made using the following: 30 μL
of 10X PCR reaction buffer, 9 μL of 50 mM magnesium chloride (New England BioLabs Inc., Ipswich, MA, USA), 1.2 μL of 25 mM dNTP’s, 2.04 μL of Taq Polymerase (Invitrogen), and 233.76 μL of water. Once the master mix was made, 26 μL was pipetted into five Eppendorf tubes. The diluted primers were then added to the five tubes. Six μL of primer pair 1 was added to the first tube; 6 μL of primer pair 2 was added to the second tube, and this proceeded for all five primer pairs. Twenty-four μL was pipetted from each of the five tubes into different wells in a 96-well PCR plate. Two rows of five wells each were used. Once the ten wells were filled with 24 μL of solution, 1 μL of cDNA was added to each well. The cDNA was selected based on the genotyping results. There were five samples that were homozygous for VIC®, and five samples that were homozygous for carboxyfluorescein (FAM). Vic® (minor allele) and FAM (major allele) are fluorescent dyes. Lastly, the 96-well plate was placed into a Veriti 96 Well Thermal Cycler (Applied Biosystems Corp., Foster City, CA, USA). The samples were run for 30 cycles at 94° C for 2 minutes, 94° C for 15 seconds, 60° C for 15 seconds, 72° C for 30 seconds, 72° C for 7 minutes, and then cooled to 25° C. Steps 2, 3 and 4 were run for 30 cycles.

Once the PCR was complete, the amplified DNA was separated through gel electrophoresis. A 20 cm X 20 cm 10% polyacrylamide gel was made using the following ingredients: 7.5 mL of 30% Acrylamide (29:1), 11.7 mL of water, 6 mL of 5X Tris-Borate-EDTA (TBE), 4.6 mL of 50% glycerol, 0.3 mL of 10% Ammonium Persulfate (APS), and 0.02 mL of TEMED. The Acrylamide was ordered from Bio-Rad Laboratories (Hercules, CA, USA), and the TEMED was ordered from Calbiochem® (San Diego, CA, USA). The gel was allowed to set and was then placed on a Vertical Gel
Electrophoresis System made by Bethesda Research Laboratories (BRL; Thermofisher, Grand Island, NY, USA). A buffer was prepared using 400 mL of 5X TBE and 1600 mL of water. The buffer was then poured into the top and bottom of the electrophoresis system. Five μL of loading dye was placed into each of the wells on the PCR well-plate. A syringe (Hamilton Co., Reno, NV, USA) was used to pipet 10 μL of a 1 Kb+ ladder (ThermoFisher, Grand Island, NY, USA) into the first well, followed by 15 μL in each of the other wells. The samples were then run on the gel using 200 volts from a Fisher Biotech Electrophoresis Systems amplifier until the DNA was far enough down the gel for distinct band resolution. Once the samples migrated far enough, the gel was removed and placed into a solution containing 5X TBE buffer and SYBR® Gold stain (Molecular Probes Inc., Eugene, Oregon, USA). SYBR® Gold is a nucleic acid gel stain. The gel was placed on a shaker for 15 minutes and then placed on a Fujifilm FLA-2000 transilluminator to view the bands made by the DNA. The software used to start, view, and manipulate the gel was Image Reader FLA-2000 (version 1.8) and Multi Gauge (version 3.0).

After testing the primers and viewing the corresponding gel, another PCR reaction was run. Five samples of homozygous VIC and 5 samples of homozygous FAM cDNA were used in the PCR reaction. The PCR product was then run on a 10% polyacrylamide gel, following the same procedure as before. All other primers were also used to amplify their respective amplicons using PCR and run on the gel. After viewing the gel in the FLA-2000, amplicons were cut for primer set 6 and primer set 4 reactions. Primer set 6 (6F-6R) amplifies an MS4A6A fragment found in all transcripts, while primer set 4 (4F-4R) amplifies a fragment that contains a retained intron 2. (Figure 1).
Purifying PCR Product

Once the bands of DNA were excised from the gels, the gel fragments were crushed and the DNA purified. DNA was purified from the gel using a QIAquick® PCR Purification Kit (Qiagen Inc., Valencia, CA, USA). This process was carried out using the protocol found for the QIAquick® PCR Purification Kit found on the Qiagen website.

After purification was complete, the samples were subjected to UV spectroscopy to quantify the DNA concentration in solution. This was done using a NanoDrop 2000 UV-Vis Spectrophotometer. First, 1 μL of water was used as a blank in the machine. Once the blank was read and the machine properly zeroed, 1 μL of the purified PCR product was placed on the reading eye. Graphs were obtained for both the primer set 6 PCR product (all transcripts) and the primer set 4 (variant 4) PCR product (data not shown).

Real-Time PCR

Quantitative real-time PCR, or qPCR, was used to monitor PCR reaction efficiency and the absolute quantification of two MS4A6A isoforms, isoform 4 and isoform 6 (whole transcript). A master mix containing 20 ng of the cDNA subjects, 10 μL of SYBR green reagent (Quanta Biosciences, Gaithersburg, Maryland, USA), 10 μL of water, and 10 μL of forward and reverse primers was used. Once the master mix is made, it was pipetted into wells, and then placed into an MJ Opticon 4 thermal cycler (Biorad). The samples were cycled under the following conditions: 95° for 3 minutes to denature, followed by 40 cycles of a 95° C denaturation step for 15 seconds and a 60° C
annealing/extension step for 45 seconds. Finally, a melting curve was used to view the amplification fidelity in combination with inspection of PCR products on an 8% TBE-PAGE gel (Simmons et al. 2011).

Using the PCR products created from this process, a standard curve was produced consisting of 10-fold serial dilutions that were amplified with the samples in each PCR run. By looking at these standard curves, PCR reaction efficiency can be monitored, and quantification of MS4A6A isoforms can occur. (Simmons et al. 2011).

Statistical Analysis

Among the cohort of different brain samples, one outlier was excluded from the statistical analysis. This sample had normalized MS4A6A isoform 4 expression much higher than the median. A linear regression analysis using SPSS software was used to analyze the association between normalized MS4A6A isoform 4 and normalized isoform 6 (whole transcript) and rs610932 genotype, AD status and microglia gene expression. Microglia gene expression, AD status and rs610932 were the independent variables.

VIII. Results and Discussion

In this experiment, we chose to look at two of the isoforms for the MS4A6A gene, isoform 6 (whole transcript) and isoform 4 (short transcript). In the statistical analyses, we chose to analyze the normalized expression of the two transcripts in association with normalized microglia gene expression. Microglia gene expression was chosen due to the previously found linkage between expression of microglia genes and the MS4A6A transcripts (Zhang et al. 2014). The plots showing the relationship between MS4A6A
isoforms and microglial content in the samples are shown in Figure 4. In Figure 4 (A), we analyzed MS4A6A Isoform 4 expression in association with microglia gene expression and the rs610932 genotype. In Figure 4 (B), we analyzed Isoform 4 expression in association with microglia gene expression and AD status. We did not find a significant association (p < 0.05) between MS4A6A Isoform 4 expression and the rs610932 genotype (p-value = .132, power = .411). We did, however, find significant associations between Isoform 4 expression and microglia gene expression (p-value = 1.1X10^{-4}, power = .986) and AD status (p-value = .002, power = .892). Microglia gene expression and Isoform 4 expression were expected to be associated due to a prior report than MS4A6A was expressed in microglia (Zhang et al. 2014).

Figure 4 (C-D) shows the analysis results for MS4A6A Isoform 6 (total MS4A6A expression). In Figure 4 (C), we analyzed Isoform 6 expression in association with microglia gene expression and the rs610932 genotype. In Figure 4 (D), we analyzed Isoform 6 expression in association with microglia gene expression and AD status. We found no significant association between MS4A6A Isoform 6 expression and the rs610932 genotype (p-value = .193, power = .339). We also found no significant association between Isoform 6 expression and AD status (p-value = .154, power = .294). We did, however, find significance between Isoform 6 expression and microglia gene expression, much like Isoform 4 (p-value = 1.0X10^{-6}, power = 1.00)

Our results show an increased expression of the short form of MS4A6A, Isoform 4, in subjects with AD. We did not, however, find a significant association between the rs610932 SNP and either Isoform 4 or Isoform 6. The latter findings are in contrast to recent findings that showed a SNP association in whole blood (Petroula et al., 2014). Our
findings also showed significant association between Isoform 4 and Isoform 6 expression and microglial gene expression, which is consistent with recent studies (Zhang et al. 2014).

These results indicate that an increase in expression of the truncated MS4A6A variant (Isoform 4), through some kind of genetic influence or other factors, is correlated with an increase in Alzheimer’s Disease prevalence. To know exactly how an increased expression of Isoform 4 alters AD likelihood is challenging. The effect size of the MS4A6A Isoform 4, which is associated with AD, is small (Petroula et al., 2014). With this in mind, Isoform 4’s impact in AD pathology risk could be represented by the differential expression that we found. The altered gene expression in AD subjects may also be a consequence of other factors. Further research is needed to evaluate the effects of over-expression of Isoform 4 in AD subjects.

With these results, MS4A6A continues to be a promising subject in Alzheimer’s research. Further research needs to be done to determine the true reason why the rs610932 SNP is correlated with increased AD risk. In addition to investigating the short form of MS4A6A’s role in AD pathology, more research needs to be done to analyze the association between the other two MS4A6A variants and AD susceptibility.

**X. Conclusions**

At the beginning of this study, we hypothesized that an increase in expression of MS4A6A Isoform 4 and Isoform 6 are correlated with an increased risk in Alzheimer’s Disease, particularly as a result of rs610932. The results of this study indicate a significant association between Isoform 4 expression and AD status. With these data in
hand, we can accept the experimental hypothesis that *MS4A6A* variant expression is associated with AD for Isoform 4. Since Isoform 6 did not have a significant association between expression and AD status, we fail to accept that experimental hypothesis. We also fail to accept the experimental hypothesis that the rs610932 SNP caused a significant association between *MS4A6A* isoform expression and AD susceptibility.

The *MS4A6A* gene is a growing subject for researchers in the Alzheimer’s field, as it has increasing potential for answers. With more and more data on this gene being produced, we are gaining more knowledge on how it can contribute to AD pathology and the possible mechanisms it uses to do so.

IX. Acknowledgements

We thank Dr. Steven Estus for his commitment and generosity in completing this experiment. We also thank Jim Simpson and Jared Vasquez for their assistance and time throughout the process. Lastly, we thank Sanders-Brown Center on Aging for permitting the use of their facilities and equipment.
## XI. Tables

### Table 1

Differential expression of *MS4A6A* in blood of control, MCI, and AD subject groups

(Petroula et al. 2014)

<table>
<thead>
<tr>
<th>MS4A6A transcript</th>
<th>Cohort (n)</th>
<th>Control – MCI</th>
<th>Control – AD</th>
<th>MCI – AD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>p Value</td>
<td>β (SE)</td>
<td>p Value</td>
<td>β (SE)</td>
</tr>
<tr>
<td>V1 MS4A6A</td>
<td>Discovery (n = 207)</td>
<td>1.88 × 10⁻⁶</td>
<td>– 0.593 (0.12)</td>
<td>0.244</td>
</tr>
<tr>
<td></td>
<td>Validation (n = 299)</td>
<td>0.854</td>
<td>0.022 (0.12)</td>
<td><strong>0.012</strong></td>
</tr>
<tr>
<td></td>
<td>Combined group (n = 506)</td>
<td><strong>0.002</strong></td>
<td>– 0.264 (0.09)</td>
<td><strong>0.009</strong></td>
</tr>
<tr>
<td>V4 MS4A6A</td>
<td>Discovery (n = 209)</td>
<td>0.107</td>
<td>0.131 (0.08)</td>
<td><strong>1.36 × 10⁻⁶</strong></td>
</tr>
<tr>
<td></td>
<td>Validation (n = 309)</td>
<td><strong>1.89 × 10⁻⁴</strong></td>
<td><strong>0.303 (0.08)</strong></td>
<td><strong>7.51 × 10⁻⁸</strong></td>
</tr>
<tr>
<td></td>
<td>Combined group (n = 518)</td>
<td><strong>1.82 × 10⁻⁴</strong></td>
<td><strong>0.213 (0.06)</strong></td>
<td><strong>9.08 × 10⁻¹³</strong></td>
</tr>
</tbody>
</table>

Total and individual variant expression was determined using real-time polymerase chain reaction. The discovery groups are subjects who had previously been analyzed by microarray, while the validation group had not been previously analyzed by microarray. Linear regression analyses were altered for age at baseline, sex, number of APOE ε4 alleles, RNA integrity number (RIN), and center. B-coefficients display the change in *MS4A6A* gene expression between the groups using the first group as baseline. Key: AD= Alzheimer’s Disease, MCI= mild cognitive impairment, SE = standard error.
XII. Figures

Fig. 1. Schematic depiction of the MS4A6A gene designating different transcript variants, location of primers, and different single nucleotide polymorphisms (SNPs) associated with the gene. V1 corresponds to Isoform 6 (whole transcript) and V4 corresponds to Isoform 4 (short transcript) (Petroula et al. 2014).

Fig. 2. Schematic representation of the human brain with a clear indication of where the anterior cingulate is located.
Fig. 3. Results of genotyping assays. The upper-left group represents samples that were homozygous for the minor allele (T). The lower-right group represents samples that were homozygous for the major allele (G). The middle group represents samples that were heterozygous for both alleles. The lower-left group represents negative controls in which no cDNA was added. VIC and FAM were used to designate major and minor alleles.
Figure 4. Isoform 4 and Isoform 6 (full-length) MS4A6A expression in brain associated with rs610932 genotype and AD status. (A) Linear regression analysis of normalized MS4A6A Isoform 4 expression revealed a significant model (adjusted R squared=.387) wherein Isoform 4 expression was associated with normalized microglial gene expression between subjects but not rs610932 genotype. (B) Linear regression analysis of normalized MS4A6A Isoform 4 expression revealed a significant model when Isoform 4 expression was associated with normalized microglial gene expression and AD status. (C) Linear regression analysis of normalized MS4A6A Isoform 6 expression revealed a significant model (R squared=.427) when the whole transcript was associated with normalized microglial gene expression between subjects but not rs610932 genotype. (D) Linear regression analysis of normalized MS4A6A Isoform 6 expression revealed a significant model when the transcript was associated with normalized microglial gene expression but not AD status.


