Title
Genetic Contributions to Alzheimer's Disease: The Role of Immune Modulatory Regions

Author
Carr, Jessie

Publication Date
2017

Peer reviewed|Thesis/dissertation

Genetic Contributions to Alzheimer's Disease: The Role of Immune Modulatory Regions

by

Jessie S Carr

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Pharmaceutical Sciences and Pharmacogenomics

in the

GRADUATE DIVISION

of the

[redacted]

Dedication and Acknowledgments

First, I'd like to thank my primary mentor during my time at UCSF, Yadong Huang. Yadong gave me the chance to work on a new and very exciting project testing a potential therapeutic candidate for Alzheimer's disease, and helped teach me the skills and approaches necessary for preclinical drug development. Yadong was my first exposure to the field of neuroscience, and it was his engaging presentation on the role of *APOE-ε4* in Alzheimer's disease that spurred my transition from cancer research to neurodegenerative diseases. Thank you, Yadong, for not only exposing me to this area of research, but for being willing to take on a first-year graduate student with absolutely no neuroscience experience and encouraging my growth and development as a scientist throughout my time in graduate school. I learned an enormous amount during my time in the Huang lab and I especially appreciate Yadong's continued support even as I transitioned to a project outside the scope of the Huang lab's main research.

I'd also like to thank Jennifer Yokoyama, my soon-to-be postdoc mentor, for enthusiastically embracing a collaboration last fall despite my lack of computational experience and very short timeline to graduation. Though it's been less than 11 months, I've gained an entirely new skillset in our time working together, from working with funnily-named R programs to starting to understand what a human brain is actually supposed to look like. I appreciate your seemingly limitless patience with me as I learn, and your constant encouragement and support of both my current work and my future goals.

I would never have decided to come to graduate school were it not for the guidance and influence of a number of people. In high school, Hank Greely and Mary

Maxon both helped open my eyes to what an exciting and innovative world scientific research was and how it could improve our society. My high school math teacher, Mr. Lowell, and my biology teacher, Mr. Domizio, both fostered my interest in their respective subjects and encouraged me to pursue a biomedical engineering degree in college. During my college career, Yu-Wang Liu, my mentor as a summer intern at OncoMed Pharmaceuticals, gave me my first exposure to actual bench research, and without her influence I don't think I would have ever considered getting a Ph.D. Yu-Wang was an excellent teacher who showed me a range of research techniques and encouraged my burgeoning curiosity about translational science. Beyond Yu-Wang, the entire OncoMed team made me excited about the field of translational research, even if I spent my first ever afternoon in the lab getting pulverized mouse organs in my hair due to exploding frozen sample tubes. I never would have gotten that internship were it not for the help and connections of many family and friends, so thank you to my mom, to Vaciliki and Stephanos Papdemetriou, and to Theo Kotseroglou for getting my resume to Tim Hoey (Go Blue!). And thank you to Chris Garcia for *not* hiring me as an intern that summer, as I know now in retrospect that my scientific passions are not in the realms of crystallography and protein structure, and I'm not sure if I would have gone to graduate school if my first exposure to science hadn't been such a good fit for me as OncoMed was.

I also owe innumerable thanks to Matt Sikora, my graduate student mentor at University of Michigan in Jimmy Rae's lab, and to the ASPET fellowship that funded my first experience in an academic lab. Having now been a graduate student mentor to an undergraduate student myself, I recognize how difficult it is to be an effective teacher

and mentor, and Matt was the best mentor I could have asked for. He not only taught me the technical skills and techniques to be successful in my project, but how to conceptualize experiments and ask interesting questions (and to write my own recommendation letters). He asked critical and insightful questions while always encouraging me, and taught me to truly think like a scientist. Matt also sparked in me an interest in pharmacogenomics, without which I'm not sure I would have ended up at UCSF for graduate school (thanks Matt!).

A variety of additional mentors at UCSF provided invaluable support and direction throughout my time in graduate school. My committee members, Jason Gestwicki and Ben Cheyette, provided critical guidance and suggestions for my project, even as I shifted from testing a preclinical therapeutic compound in mouse models to human genetic disease association studies, and were incredibly supportive with my non-conventional path to graduation over the past year. Lastly, Deanna Kroetz, the director of the PSPG graduate program, helped me immeasurably throughout my career as a graduate student and was never too busy to answer a panicked email or have a late-night meeting to help me solve a problem.

Thank you to the Huang lab for four years of learning, growth, and friendship, especially to the scientists, postdocs and older graduate students who taught me techniques, encouraged me, and helped guide my development as a scientist. I'd like to thank the members of the Yokoyama Lab (aka Laboyama) for welcoming me with open arms, and I'd especially like to thank Natasha Steele for embracing me on the HLA project with such enthusiasm and warmth (and such good BIGDAWG tutorials!). Beyond my lab communities, I'm grateful to have been a part of the PSPG community. I know

my time at UCSF would have been entirely different had I been in a different graduate program, and I'm happy and proud to be part of the PSPG family.

While I don't know them personally, I'd also like to recognize all the participants at UCSF and beyond who choose to participate in clinical research – without their participation, none of my work would be possible. I hope my work can contribute to the Alzheimer's patient community in return for all the participants and families have contributed to science.

Lastly, I would like to thank all my friends and family who supported me through the process of getting my Ph.D., especially my parents, Marion and Scott. My mom listened to me talk about the ups and downs of science for probably thousands of hours on my daily walk home, and both of them provided critical outside perspective during my time at UCSF (even if I didn't always listen). Thank you both for always making it possible for me to pursue my goals of a career in science, whether through sub-market rent or your continuous support and encouragement.

Contributions:

Chapter I is in preparation for submission to a book titled "Genetics and Epigenetics through Development" (editors Christopher Bartlett and Stephen Petrill, Ohio State University). Text was written by Jessie Carr with editing and content suggestions from both Ethan Geier and Jennifer Yokoyama.

Chapter II contains text and figures from a previously published manuscript:

Steele NZR, Carr JS*, Bonham L*, Geier E, Damotte V, Miller ZA, Desikan RH, Boehme KL, Mukherjee S, Crane PK, Kauwe JSK, Kramer JH, Miller BL, Coppola G, Hollenbach JA, Huang Y, Yokoyama JS. Fine-mapping of the human leukocyte antigen (HLA) locus as a risk factor for Alzheimer's disease: A Case-Control Study. PLoS Medicine. 2017;14(3): e1002272. doi:10.1371/journal.pmed.1002272.

Jessie Carr ran all analyses involving the ADGC dataset (n = 11,381) and led revision of the manuscript after the initial submission. JS Carr performed quality control on the ADGC dataset (with help from E Geier), imputed all the HLA alleles, and ran all the primary and secondary HLA statistical analyses for the ADGC cohort. JS Carr performed a merged MAC + ADGC analysis, a specific test of interaction between HLA haplotypes and *APOE-ε4*, and sex-specific analyses, at both the haplotype and allelic level. JS Carr also performed an iterative analysis of the full cohort divided into smaller cohorts to help alleviate concern over possible Type I error. NZR Steele and JS Yokoyama conceived of the original experiments. JS Yokoyama and L Bonham performed the analyses of haplotype associations with clinical measures. KL Boehme, S Mukherjee, PK Crane and JSK Kauwe provided great assistance through their preparation of the merged ADGC dataset allowing us to use individuals from over 30 publicly available studies.

Chapter III is in preparation for submission. Jessie Carr performed all quality control and all analyses. JS Carr and JS Yokoyama conceived of and designed all experiments in this study.

Genetic Contributions to Alzheimer's Disease:
The Role of Immune Modulatory Regions

Jessie S Carr

Abstract

Alzheimer's disease (AD) is a progressive disorder that affects cognitive function. There is increasing support for the role of neuroinflammation and aberrant immune regulation in the pathophysiology of AD. The role of the Human Leukocyte Antigen (HLA) region in neurodegenerative disease is becoming increasingly appreciated as numerous genome-wide association studies (GWAS) identify significant associations with this complex genomic region and disease risk.

We used a robust imputation method on two case–control cohorts (a small UCSF cohort and a large cohort from the Alzheimer's Disease Genetics Consortium [ADGC]) to identify HLA haplotypes associated with Alzheimer's disease and followed up these studies with direct sequencing of the HLA region in AD cases and controls, including both typical amnestic and atypical clinical forms of disease. In our imputed study, we found the haplotype *A*03:01~B*07:02~DRB1*15:01~DQA1*01:02~DQB1*06:02* (p = 9.6×10^{-4}, odds ratio [OR] [95% confidence interval] = 1.21 [1.08–1.37]) was associated with increased risk of AD in the combined UCSF + ADGC cohort (n = 11,690). Secondary analysis suggested that this effect may be driven primarily by individuals who are negative for *APOE-ε4*. Separate analyses of class I and II haplotypes further supported the role of class I haplotype *A*03:01~B*07:02* (p = 0.03, OR = 1.11 [1.01–1.23]) and class II haplotype *DRB1*15:01~DQA1*01:02~DQB1*06:02* (*DR15*) (p = 0.03, OR = 1.08 [1.01–1.15]) as risk factors for AD. We followed up these genetic

associations in a separate clinical dataset representing the spectrum of cognitively normal controls, individuals with mild cognitive impairment, and individuals with AD to assess their relevance to disease. Carrying *A*03:01~B*07:02* was associated with higher CSF amyloid levels. We also found a dose-dependent association between the *DR15* haplotype and greater rates of cognitive decline on two different assessments.

We also directly sequenced the HLA region in an expanded cohort of AD cases, controls, and atypical AD cases seen at UCSF. We corroborated the accuracy of HLA imputation via direct sequencing of 308 overlapping samples and confirmed the association of the haplotype previously associated with AD risk in the UCSF cohort. We also found that the *A*03:01~B*07:02~C*07:02~DRB1*15:01~DQB1*06:02~DPB1*04:01* haplotype was associated with decreased risk of atypical AD in our cohort (p = 0.01, OR = 0.18 [0.02-0.74]). Taken together, our findings corroborate a role of the HLA in AD risk and suggest a differential role of HLA variation in amnestic versus atypical AD.

Table of Contents

List of Figures

List of Tables

Chapter I:

Genetics of neurodegenerative diseases: An overview

Neurodegenerative diseases are characterized by the progressive degeneration of the structure and function of specific components within the nervous system, leading to changes in cognition, behavior, and/or movement, and ultimately resulting in death. The neurodegenerative diseases described in this chapter, including Huntington's disease (HD), Parkinson's disease (PD), amyotrophic lateral sclerosis (ALS), frontotemporal lobar degeneration (FTLD), and Alzheimer's disease (AD), are devastating diseases that affect millions of people. There are currently few effective therapies for this class of disorders. In general, neurodegenerative diseases often present with similar or overlapping clinical features, and may share underlying pathological features resulting from specific forms of protein misfolding and aggregation, alterations in neuronal function, and mitochondrial dysfunction. While the contributions to neurodegenerative disease risk are generally thought to be multifactorial in nature, all of these diseases have a strong heritable component, with many risk-associated loci shared across multiple neurodegenerative diseases [1–3]. In addition, a subset of neurodegenerative diseases is familial, most often inherited in an autosomal dominant pattern.

Identification of genetic risk factors for neurodegenerative diseases began in the late 1980s and early 1990s with linkage analysis studies in families with a history of inherited neurodegenerative disease. These studies identified large genomic regions within a given family pedigree that segregated with disease carriers, and required subsequent fine-mapping studies to identify the causative gene and/or specific genetic variants underlying disease. These early studies commonly identified rare variants that often only account for a very small proportion of all disease cases. The advent of genome-wide association studies (GWAS) in the mid-2000s allowed researchers to

assess the association with disease risk of millions of common single nucleotide polymorphisms (SNPs) that occur across the genome at a frequency of ≥ 1%. GWAS have been reasonably successful, identifying a number of genetic loci associated with risk for different neurodegenerative diseases. However, these studies require large cohorts of patients since common variants typically have modest risk contributions to disease. Finally, candidate gene studies examine the association between specific genetic variation and disease risk based on evidence that a gene of interest is related to the disease through underlying biological function. Findings from these three types of studies have contributed to our understanding of how genetic variation contributes risk to development of neurodegenerative disease, and to the underlying etiology of these diseases.

This chapter presents a basic overview of the most important genetic risk factors for five different neurodegenerative diseases. The lists of genes and variants associated with each disease are not intended to be comprehensive, particularly with respect to common genetic variants identified by GWAS, as that is beyond the scope of this chapter. Rather, we use the identified genetic loci as a means of highlighting how genetic discoveries — particularly in familial forms of disease—inform our understanding of underlying disease pathobiology and can provide candidates for therapeutic intervention. While findings from genetic studies have greatly enhanced our understanding of the biological processes underlying these diseases, it should be noted that our understanding of the causes of these diseases — both genetic and non-genetic — is constantly evolving. The increasing availability of deep sequencing information and detailed, multi-modal phenotypic data from diverse populations are some of the

mechanisms through which future studies may elucidate additional genetic contributions to these devastating diseases. Moreover, much work remains to be done to identify effective disease-modifying treatments and cures for neurodegeneration.

Huntington's disease

Also known as Huntington chorea, HD is a dominantly transmitted, progressive neurodegenerative disorder that leads to cognitive impairments, motor abnormalities, and psychiatric disturbances [4]. In populations of European ancestry, the prevalence of HD is approximately 5-7 affected individuals per 100,000 [5]. Symptoms typically occur starting around 35-44 years of age, and affected individuals survive an average of 15-18 years after symptom onset [4].

HD: Trinucleotide CAG repeat expansions in HTT

Early studies in the 1980s linked HD to chromosome 4 [6], though the exact gene and genetic alteration responsible for the disease remained a mystery until the early 1990s. In 1993, studies identified the Huntingtin gene (*HTT*) and discovered that longer expansions of the repetitive trinucleotide sequence CAG within the *HTT* gene appeared in families with disease, with affected individuals carrying 42-100 CAG repeats [7].

Further research has elucidated a length-dependent effect of the CAG repeat expansion, with increased expansion size associated with greater risk of developing symptoms and transmitting disease to offspring. The median CAG repeat length observed in unaffected individuals is 18, and up to 26 repeats does not confer any risk for developing or transmitting HD [8] (Table 1). Individuals carrying "intermediate alleles" or "mutable normal alleles" (27-35 CAG repeats) will not develop HD, but the meiotic

instability of these repeats means that these individuals may transmit a slightly smaller or slightly larger repeat to their offspring [9–11]. While carrying a repeat of this size is relatively rare (approximately 3.2% of all repeats are in this range) [12], if the repeat expands over the threshold of 36 CAG repeats, the offspring of an unaffected parent will be at risk for HD. The risk of CAG repeat expansion is associated with both the sex and the age of the transmitting parent, with older parents and males being more likely to transmit an expansion [9,13,14]. HD risk associated with the 36-39 CAG repeat range is less clear. No individuals with fewer than 36 CAG repeats have ever been diagnosed with HD [9], but some individuals who carry 36-39 CAG repeats develop symptoms of HD while others do not, suggesting that these repeat lengths are not fully penetrant [15–18]. Individuals with 40 or more CAG repeats will develop HD and have a 50% chance of transmitting the disease-causing allele to an offspring.

HTT expansion length is also associated with age of symptom onset. Greater numbers of CAG repeats are associated with earlier disease onset [15,19,20], and individuals with 60 or more CAG repeats develop juvenile HD, which has an age of onset as early as 3 years old [21]. Approximately 60% of the variance in HD age of onset can be explained by CAG repeat length, while additional modifying genes, as described below, and environmental factors likely contribute to the remaining phenotypic variance [5]. An excellent historical perspective of the molecular genetics of HD is provided in Gillian Bates' review [22].

Table 1. CAG repeat expansion length in Huntington's disease

Number of Repeats	Classification	Effect
≤ 26	Normal	Individual not at risk for HD
27 – 35	Intermediate	Individual not at risk for HD, but has risk of passing allele in disease-causing range to offspring
36 – 39	Reduced-penetrance HD	Individual may or may not develop HD, has 50% risk of passing disease allele to offspring
≥ 40	Full-penetrance HD	Individual will get HD and has 50% risk of passing disease allele to offspring

More recently, candidate gene studies and GWAS have aimed to identify genetic variation outside of *HTT* that contributes to clinical features of HD, including age of onset. While studies of the *HTT* gene itself and other candidate genes have failed to provide convincing and replicable evidence of genetic modifiers for HD risk [23], a recent GWAS of over 4,000 HD patients identified several loci that do not contribute to HD risk by themselves but, in combination with a sufficient CAG repeat length, can modify age of disease onset. Two variants within a locus on chromosome 15 were associated with either 6 years earlier onset or 1.4 years later onset, while a locus on chromosome 8 was associated with disease onset 1.6 years earlier. Pathway analysis in this study suggests that DNA handling and repair mechanisms may be responsible for modifying age of symptom onset in HD [24].

Mapping of HTT repeat expansion highlights importance of huntingtin in disease pathogenesis

Within the context of the neurodegenerative diseases described in this chapter, HD is unique in that a single pathogenic variant has been established to cause this disease.

Based on identification of CAG repeats in *HTT* as causative of HD, researchers have focused on understanding the normal and pathologic role of the *HTT* protein product, huntingtin, with a particular focus on the brain regions and neuronal subtypes most affected by disease. The CAG expansion is believed to cause neurodegeneration through the accumulation of mutant huntingtin and polyglutamine (polyQ) aggregates at the protein level, rather than the effects of the CAG expansion at the DNA or mRNA level [25]. Larger CAG repeat expansions are thought to confer a toxic gain of function to the huntingtin protein as polyQ repeats aggregate in the brains of HD patients and lead to a host of deleterious effects. These polyQ expansions can form amyloid-like aggregations *in vitro* and *in vivo* [26] and are thought to be involved in transcriptional dysregulation, deficits in protein folding and degradation, impaired energy metabolism and mitochondrial dysfunction, and perturbations in neuronal circuitry [27].

The molecular underpinnings of HD and the progressive deterioration of executive function due to the loss of neurons that is correlated with the accumulation of pathogenic protein aggregates is strikingly similar to other neurodegenerative disorders. Additionally, the trinucleotide repeat expansion in *HTT* observed in HD is just one disease within an entire class of trinucleotide repeat disorders, including spinal and bulbar muscular atrophy and many types of spinocerebellar ataxia. This nucleotide repeat expansion, while different in exact nucleic acid sequence and repeat length, is also comparable to the hexanucleotide repeat expansion in *C9ORF72*, which is the most common genetic cause of the ALS-FTLD spectrum of disorders [28].

HTT and therapeutic strategies in Huntington's disease

HD provides a sobering lesson on the immense difficulties of developing therapies for neurodegenerative diseases. HD is unique amongst neurodegenerative diseases in that a single pathogenic variant is solely responsible for disease risk, with essentially full penetrance. For many neurodegenerative diseases, a complete understanding of the genetic contributions to disease risk is seen as the holy grail that will enable the field to fully understand disease pathogenesis, and thus identify targets for intervention and develop effective therapies. However, despite our knowledge — for over 20 years — of this singular genetic variant that leads to HD, we still lack any effective therapies for this devastating disease. One exciting avenue of therapeutic intervention in HD involves directly targeting the pathogenic *HTT* RNA through RNA interference or antisense oligonucleotides — while these therapeutic approaches are still in development, this strategy of reducing RNA has been approved by the FDA for the treatment of other diseases and may also be useful in other neurodegenerative diseases such as AD, the ALS-FTLD spectrum, or tauopathies (diseases characterized by the pathological aggregation of misfolded tau protein).

Parkinson's disease

Worldwide, PD is the most common movement disorder, and second most common neurodegenerative disease after AD [29]. PD is clinically characterized by the presence of four typical motor symptoms (or a subset of these symptoms): bradykinesia (slowness of movement), rigidity, resting tremor, and/or postural instability [30]. Definitive diagnosis of PD requires neuropathological examination of the substantia nigra pars compacta to confirm the loss of dopaminergic neurons within this region. In

the remaining neurons, the neuropathological hallmark of PD is the presence of aggregated α-synuclein protein in Lewy bodies or Lewy neurites. Although this 'typical' neuropathological hallmark is observed in the majority of PD cases, it is missing in a subset of cases [29].

PD was considered a prototypical non-genetic disorder, potentially caused by environmental factors such as viral illness or exposure to specific chemicals, until 1997 when the first causal pathogenic variant for PD was discovered [31]. Subsequent studies have used a variety of techniques, including linkage analysis, direct sequencing, and genetic association studies, to identify further genetic variants that contribute to PD risk. GWAS in particular have identified 26 risk loci for PD, suggesting the 24 genes nearest to these SNPs may be involved in PD risk [29]. Additional chromosomal regions have been linked to PD risk, though the exact genes responsible for disease risk within these regions have yet to be determined [32].

Autosomal dominant PD: SNCA, LRRK2, and VPS35

The first gene linked to PD was an autosomal dominant A53T variant in *SNCA* (synuclein alpha) identified in 1997 in three unrelated Greek families and an Italian family with PD [33]. The *SNCA* gene encodes for the α-synuclein protein, which within three months of the discovery of *SNCA* variants was shown to be the primary component of Lewy body pathology in PD [34]. Additional variants within *SNCA* have been identified as causative for PD through family studies, including A30P [35], E46K [36], H50Q [37,38] and G51D [39,40]. Other studies have also identified structural variants of *SNCA* as PD risk factors, such as gene duplications [41,42] and triplications [43]. Interestingly, triplications of *SNCA* lead to an earlier age of onset by ~10 years

compared to duplications [44]. Similar to when *APP* and *PSEN1/2* pathogenic variants were found to cause familial autosomal dominant AD (described in detail later in this chapter), the discovery of *SNCA* variants led to increased research focus on α-synuclein in PD pathogenesis. Even though these pathogenic SNPs and structural variants in *SNCA* are rare and only account for a very small proportion of PD cases, more recent GWAS have identified common variants within *SNCA* that are also associated with PD risk [45–47]. Beyond the presence of accumulated α-synuclein in the neuropathological hallmark of PD (Lewy bodies), the α-synuclein protein is also involved in assembling a protein complex that is involved in neurotransmitter release from the presynaptic terminal [48], implicating the process of synaptic transmission in the etiology of PD.

In 2004, autosomal dominant pathogenic variants in *LRRK2* (Leucine-rich repeat kinase 2, initially known as *PARK8*) were found to segregate with PD cases [49,50], and subsequent studies revealed seven highly penetrant pathogenic variants in *LRRK2*: G2019S, R1441G/C/H, I2020T, Y1699C and N1437H [51,52]. The most common of these variants, G2019S, occurs in 1% of sporadic PD patients and 4% of familial PD cases, and increases the risk of developing PD by 74% by age 79. The frequencies of each *LRRK2* variant differ across ancestral populations [51], and different *LRRK2* mutations are responsible for varying incidences of disease in different populations; for example, the G2019S mutation is responsible for 1-7% of PD patients with European ancestry, 20% of cases in Ashkenazi Jewish PD patients, and up to 40% of PD cases in patients of Arab ancestry [51]. Although LRRK2's function is currently unknown, the protein has kinase and GTPase domains, and has been implicated in cellular processes involved in other neurodegenerative diseases, including vesicle trafficking and

autophagy [53]. Just as GWAS identified common variations in *SNCA* as PD risk factors, common variants in *LRRK2* have also been implicated as risk factors for PD [54,55], further supporting a role of the LRRK2 protein in disease pathogenesis.

More recently, whole exome sequencing of a Swiss family and an Austrian family with late-onset, autosomal dominant PD identified variants in *VPS35* (vacuolar protein sorting 35) segregating with affected individuals in these families [56,57]. *VPS35* encodes an essential component in the retromer complex, which is important in protein transport and recycling pathways [58]. As retromer dysfunction has also been implicated in AD, the identification of variants in *VPS35* as risk factors for PD further highlights how disparate neurodegenerative diseases may have similar underlying pathological processes [59].

Recessive PD: PARK2 (Parkin), PINK1, PARK7 (DJ-1), and DNAJC6

To date, pathogenic variants in four genes have been established to cause PD in an autosomal recessive mode of inheritance. Although these variants are rare, they are enriched in early-onset PD cases (< 50 years). Almost one-third of PD patients under the age of 30 years old carry recessive variants in one of these four genes [60]. The first gene in which recessive pathogenic variants were found to cause PD was *PARK2/Parkin*, an ubiquitin protein ligase involved in targeting proteins for degradation [61]. Variants in *PARK2* are implicated in approximately 8% of early-onset PD cases [60]. Subsequent studies of families with early-onset PD identified recessive variants in *PINK1* (PTEN-induced putative kinase 1) as the next most common genetic cause of early-onset PD, explaining approximately 4% of early-onset cases [60,62]. Both *PINK1* and *PARK2* have been implicated in mitophagy, the process of degrading mitochondria

via autophagy [63], suggesting a role for this process in the pathogenesis of PD. Recessive variants in *DJ-1/PARK7*, a component in the cell's response to oxidative stress, were first associated with PD in 2003 [64], but are considered an extremely rare cause of PD and account for less than 0.5% of early-onset cases [60]. Interestingly, in tauopathies, DJ-1 protein co-localizes with tau inclusions [65]. DJ-1 is also hypothesized to provide functional relevance parallel to the *PINK1/PARK2* mitophagy pathway in response to oxidative stress [66], again highlighing a potential connection between distinct neurodegenerative diseases and their underlying pathogenic processes. Most recently, variants in *DNAJC6* (DNaJ heat shock protein family 40 member C6; auxilin) were found to cause both juvenile Parkinsonism, in which symptoms develop in patients younger than 11 years old [67,68], and early-onset PD [69]. Auxilin protein is involved in endocytic/lysosomal trafficking [67], yet again emphasizing the importance of protein homeostasis as a shared pathogenic process between multiple neurodegenerative diseases.

Non-Mendelian PD: HLA-DQB1, MAPT, and other loci

A recent meta-analysis of GWAS for PD across 19,061 PD cases and 100,833 controls identified 28 independent risk variants for PD [45]. The effect size of each individual variant is small, but cumulatively these loci may account for a substantial increase in disease risk [45]. Several of the loci identified in this meta-analysis, previous GWAS, and other prior association studies are of particular relevance to this chapter, given the implication of these same loci in other neurodegenerative diseases. Interestingly, this meta-analysis and several prior studies implicated a role of the immune system — and particularly variation in the human leukocyte antigen (HLA) system — in PD risk. While

this study identified a protective SNP proxying *HLA-DQB1* that was associated with decreased risk of developing PD [45], prior studies identified SNPs proxying *HLA-DRA* [70] and *HLA-DRB1* [71] that associated with increased PD risk. Given the extreme degree of variation in the HLA region, the strong linkage disequilibrium between various HLA alleles, and the significant effect of genetic ancestry on these factors, fine-mapping the region to identify the specific loci responsible for protective and risk-associated signals is a challenging task. However, increasing evidence implicating the immune system in other neurodegenerative diseases [72] underscores the importance of understanding the contribution of this complex genetic region to the etiology of these diseases.

Another risk loci shared between PD and other neurodegenerative diseases is *MAPT* (microtubule-associated protein tau), which encodes the tau protein. Tau aggregation is the pathological hallmark of tauopathies, including progressive supranucelar palsy (PSP), corticobasal degeneration (CBD), chronic traumatic encephalopathy (CTE), and Pick's disease. Furthermore, neurofibrillary tangles of tau protein are one of the two neuropathological hallmarks of AD, along with plaques made up of aggregated amyloid beta. The *MAPT* region was found to be associated with PD risk through both GWAS and a number of candidate gene studies. Eventually an extended haplotype across a large region of chromosome 17q21 termed the H1 haplotype, which includes *MAPT,* was identified as the responsible risk locus for PD [73]. The H1 haplotype has a stronger association with PD patients who also have dementia compared to PD patients without dementia [74], suggesting tau contributes to these cognitive changes. One SNP within this region in particular has been implicated in

risk for both PD and AD [1]. While sub-haplotype variation within the H1 haplotype is subtly different between PD and other neurodegenerative diseases (for example, compared to controls the H1-rs242557^{A} sub-haplotype is increased in individuals with PSP/CBD while the H1-rs242557^{G} sub-haplotype is increased in PD) [75], the involvement of this overall genetic region in multiple neurodegenerative diseases again highlights the potential for common underlying disease mechanisms.

Genetic causes of PD converge on critical roles of α-synuclein aggregation, mitochondrial function, synaptic transmission, and endosomal/lysosomal trafficking and recycling in PD pathogenesis

Although PD was originally thought to primarily have an environmental etiology, the studies discussed in this section provided crucial evidence for the contribution of genetic risk factors to PD. Elucidation of many genetic risk factors for PD has spurred great advances in our understanding of the pathological processes underlying disease. The genes implicated in PD risk described above connect PD with numerous biological functions, including pathogenic accumulation of α-synuclein, impaired mitochondrial function, synaptic transmission, and dysregulation of endosomal/lysosomal pathways involved in protein trafficking and recycling. α-synuclein is the protein aggregate in the primary neuropathological hallmark of PD, Lewy bodies, and it is unsurprising that pathogenic variation within the *SNCA* gene can cause disease. *PARK2/parkin*, *PINK1*, and *DJ-1* are all directly involved in mitochondrial energy metabolism, while overexpression of α-synuclein or pathogenic variation within *LRRK2* have also been shown to impair mitochondrial function *in vitro* and/or *in vivo* [76]. Taken together, the involvement of numerous genes that affect mitochondrial function suggest that deficits

at several steps within this critical process may lead to PD. Beyond mitochondrial function, the genetics of PD have implicated other biological processes in the pathobiology of PD, including synaptic transmission. α-synuclein is not only the primary component of Lewy bodies but is also involved in the formation of a protein complex that plays a role in neurotransmitter release. *LRRK2* has also been shown to be involved in multiple facets of synaptic transmission [76]. Lastly, several genes implicated in PD risk (including *DNAJC6/*auxilin and *VPS35*) are involved in protein trafficking and/or recycling through endosomal/lysosomal systems, highlighting the potential involvement of these pathways in disease etiology. Despite the advances in understanding of PD in the past few decades, a large component of PD heritability remains unexplained, and continued efforts are required to provide a more complete picture of the genetic contributions to PD risk.

Identification of dopaminergic neuron loss in the substantia nigra as a hallmark of PD has resulted in development of relatively effective therapies for treating symptoms related to PD, primarily through increasing dopamine levels or mimicking the effects of dopamine in the brain. However, these therapies can only slow the clinical progression of disease. Hope remains that further understanding of the genetics that underlie PD may inform development of a true cure or disease-preventing therapeutic strategy for this common disorder.

Amyloid Lateral Sclerosis and Frontotemporal Lobar Degeneration spectrum

Amyloid lateral sclerosis (ALS, also commonly known as Lou Gehrig's disease or motor neuron disease [MND]) is a neurodegenerative disorder characterized by the degeneration of motor neurons in the cerebral cortex, brainstem, and spinal cord

leading to progressive paralysis and eventual death from respiratory failure. The progression of this devastating disease is particularly quick, with most ALS patients succumbing to disease within 2-5 years of symptom onset. While the prevalence of this disease is estimated at ~5 cases per 100,000 people [77], these numbers are skewed by the short duration between diagnosis and death. This is supported by recent estimates of ALS being implicated in more than 1 in 500 adult deaths in the US and UK [78]. Neuropathologically, brains of patients with ALS contain protein inclusions, the vast majority (97%) of which contain TDP-43, while a small percentage are composed primarily of SOD1 (2%) or FUS (<1%) [79]. Approximately 10% of ALS cases are familial, most of which are inherited in an autosomal dominant fashion, while the remaining 90% are considered to be sporadic [80].

Frontotemporal dementia (FTD) is the umbrella clinical diagnosis that represents a large and heterogeneous group of clinical syndromes due to underlying frontotemporal lobar degeneration (FTLD) pathology. FTD is the second most prevalent form of dementia after AD and affects approximately 10-30 per 100,000 people [81], with approximately 40% of all cases considered familial [82]. Clinically, there are three major subtypes of FTD: behavioral variant FTD (bvFTD), characterized by early changes to behavior and personality, and non-fluent variant primary progressive aphasia (nfvPPA) and semantic variant primary progressive aphasia (svPPA), which are characterized by speech and language impairments [83,84]. In 75-80% of FTD cases, the age of onset is < 65 years of age [81], making FTD a leading cause of presenile dementia. Similar to the clinical symptomology, the underlying neuropathological characteristics of FTLD are heterogeneous, and are generally categorized by the major

protein present in pathological inclusions. The majority of FTLD cases have either tau inclusions (45%) or TDP-43 inclusions (45%), while a smaller fraction are positive for FUS (9%), UPS (ubiquitin-proteosome system; 1%), or do not contain protein pathology [79]. However, all FTLD cases are characterized by progressive neuronal loss and/or dysfunction in the frontal and temporal lobes. FTLD diagnoses are composed of 3-repeat tau pathology (Pick's disease), 4-repeat tau pathology (with tufted astrocytes in PSP or astrocytic plaques in CBD) [85], and TDP-43 pathology (Types A, B, C, or unspecified) [86].

Historically, ALS and FTD were thought to be distinct, unrelated disorders, and the initial findings of genetic associations with these diseases (*SOD1* in 1993 for ALS, *MAPT* in 1998 for FTLD) were consistent with this hypothesis. However, recent discoveries of shared clinical symptoms, overlapping neuropathological hallmarks, and common genetic risk factors suggest that ALS and FTD exist on a spectrum, with the extreme ends of the spectrum being patients showing 'pure' ALS or 'pure' FTD and a middle ground consisting of patients with genetic and/or phenotypic characteristics of both diseases.

Clinically, a significant proportion of FTD patients display motor impairments, with ~40% of cases displaying measurable motor dysfunction (likely due to underlying PSP or CBD pathology), and 15% meeting clinical criteria for a diagnosis of ALS. Similarly, ALS often occurs in conjunction with cognitive and/or behavioral changes, with up to 50% of ALS patients displaying some symptoms falling within the FTD spectrum and ~15% of ALS patients meeting formal criteria for a clinical diagnosis of FTD [87–90]. As mentioned above, the neuropathological presence of TDP-43 in tau-negative, ubiquitin-

positive protein inclusions has been identified in both FTD and ALS [91,92]. Notably, TDP-43 pathology is not unique to the ALS-FTD disease spectrum, as it has also been observed in a substantial number of AD and dementia with Lewy bodies (DLB) patients [93]. Indeed, the preponderance of mixed pathology in neurodegeneration is underappreciated and likely more common with increasing age [94], though there are often clear primary pathological diagnosis that are most congruent with clinical diagnosis, with co-pathology modestly or minimally contributing to clinical syndrome. The genetic risk factors for ALS and FTD are discussed in more detail below, beginning with ALS, then moving to shared risk factors, and concluding with FTD.

Autosomal dominant ALS: SOD1 and other genes

The first discovered pathogenic variant causative for ALS was *SOD1* (superoxide dismutase 1), identified in 1993 [95]. To date, 185 pathogenic variants in *SOD1* have been identified [96] and these variants are responsible for ~13% of familial ALS cases and < 1% of sporadic cases [80]. Variants in *SOD1* have variable frequencies across different ethnic populations, and exert strikingly different effects on disease progression. For example, A4V is the most common pathogenic *SOD1* variant in the US (responsible for ~50% of *SOD1* ALS cases) and is associated with extremely rapid disease progression and death within one year of symptom onset, while the most common *SOD1* variant in Japan (implicated in ~40% of Japanese *SOD1* ALS cases) is H46R and is associated with an extended disease course of approximately 15 years [97,98]. Highlighting the effect of *SOD1* variation on the extreme ALS end of the ALS-FTD spectrum, cases of *SOD1*-associated ALS generally do not develop cognitive impairment [99] and are considered neuropathologically distinct from other types of ALS

due to their lack of TDP-43 and/or FUS inclusions [100]. *SOD1* variants are thought to contribute to ALS through a variety of pathways, including mitochondrial dysfunction, oxidative stress, and pathogenic aggregation of misfolded proteins [101], and therapeutic treatments targeting *SOD1* are in clinical trials [102].

In addition to *SOD1*, rare variants in several other genes have been linked to inherited forms of ALS. Moreover, common variants identified as risk factors or disease modifiers have been identified in large GWAS of ALS patients. For a detailed description of these genetic risk factors, refer to Alan Renton's review [99].

ALS-FTLD: TARDBP, FUS, C9ORF72, VCP, UBQLN2, SQSTM1

Shortly after the discovery of TDP-43 positive inclusions in the brains of both ALS and FTLD cases, follow-up studies identified pathogenic variants in the gene encoding this protein, *TARDBP* (TAR DNA binding protein), in ALS [103–106], FTD [107–109], and ALS-FTLD cases [107,110,111]. While *TARDBP* pathogenic variants occur throughout the ALS-FTLD spectrum, these variants tend to occur more frequently in ALS cases, accounting for 3-4% of familial ALS cases and approximately 1% of sporadic ALS, and occur only very rarely in pure FTLD [93]. As TDP-43 is a DNA- and RNA-binding protein that regulates transcription and splicing (among other functions), identification of *TARDBP* variants in these diseases spurred interest in the role of RNA processing in disease pathogenesis across the ALS-FTLD spectrum.

Shortly after linking *TARDBP* variants to ALS and FTLD, pathogenic variants in *FUS* (fused in sarcoma) were linked to ALS [112,113]. Like TDP-43, FUS is involved in RNA binding and processing, and ALS cases with *FUS* pathogenic variants are characterized by FUS-positive inclusions rather than TDP-43- and ubiquitin-positive

inclusions. Though only a few rare *FUS* variants of uncertain pathogenicity have been identified in clinical FTD or ALS-FTD patients [114,115], FUS-positive inclusions are present in up to 9% of FTLD patients [79]. Furthermore, there appear to be domain-specific effects of pathogenic variants in *FUS* on disease phenotype, with variants that strongly affect the nuclear localization signal of FUS associated with younger ages of disease onset [93]. Since the FUS and TDP-43 proteins have similar functions in RNA processing and regulation, and pathogenic variants in both genes are linked to ALS, it has been suggested that motor neurons are particularly sensitive to disruptions of RNA processing [116].

One of the most seminal discoveries in the ALS and FTD fields — and undoubtedly one of the most pivotal findings for the reclassification of these diseases as a spectrum instead of disparate clinical syndromes — was the finding that a repeat expansion in *C9ORF72* caused both ALS and FTLD [28,117,118]. While the repeat expansion (a repeat of six nucleotides, GGGGCC) may vary between 2-23 hexanucleotide repeats in healthy individuals, FTLD and ALS patients have dramatically higher repeat expansions, in the range of 700-1600 repeats [28]. A cut-off of >30-35 repeats has been tentatively established as the pathogenic threshold, though shorter repeat lengths have in rare cases been observed in ALS or FTD patients [119,120]. Intermediate length repeats between 23-30 have been investigated for their role as risk factors in ALS, FTD, and even PD [121], with variable findings [119,120]. While the frequency of pathogenic repeat expansions in *C9ORF72* varies by ancestral population, this genetic alteration is widely recognized as the most common cause of familial ALS (~40%) and FTLD (~25%), and underlies a significant number of sporadic cases of both

these diseases (~6%) [122]. The penetrance of pathogenic repeat expansions appears to be age-dependent, with no instances of disease in expansion carriers under 35 years of age, 50% penetrance in carriers age 58, and almost complete penetrance in carriers 80 years and older [122]. Clinically, FTLD and ALS *C9ORF72* repeat expansion carriers have considerable symptom heterogeneity [123]. Repeat expansions in *C9ORF72* have also been implicated in rare cases of PD [124] and AD [125–127], as reviewed elsewhere [128]. Many hypotheses have been suggested for the pathogenic involvement of these intronic repeat expansions in FTLD and ALS. These hypotheses include loss of function of the C9ORF72 protein, toxic accumulation of dipeptide-repeat proteins produced by repeat-associated non-ATG-dependent translation, and toxic accumulation of RNA foci leading to the sequestration of RNA-binding proteins and subsequent dysregulation of RNA processes including splicing, trafficking and translation (reviewed in [129,130]). Further understanding the molecular mechanisms by which *C9ORF72* expansions contribute to disease is an important focus of ongoing research in the ALS-FTLD field and has already resulted in preclinical studies for therapeutic intervention through anti-sense oligomers targeting the expansion [131].

Rare variants in several other genes have been implicated in risk for ALS, FTLD, and ALS-FTLD. Pathogenic variants in *VCP* (valosin containing protein) were initially identified in a rare, autosomal dominant subtype of FTLD known as inclusion body myopathy associated with Paget's disease of the bone and frontotemporal dementia (IBMPFD) [132,133]. *VCP* encodes for a protein critical in protein degradation through the ubiquitin-proteasome complex [134] and *in vitro* and *in vivo* studies have highlighted the involvement of VCP in TDP-43 pathology (reviewed in [135]). In 2010, whole exome

sequencing of several Italian families linked pathogenic variants in *VCP* to autosomal dominant ALS and ALS-FTLD [136]. Subsequent studies confirmed the link between *VCP* and ALS, and have identified pathogenic variants in *VCP* in rare cases of sporadic ALS [137–139].

Pathogenic variants in *UBQLN2* (ubiquilin 2), another gene with an important role in ubiquitin-mediated protein degradation, were linked to ALS and ALS-FTLD in 2011 [140], and later with rare cases of pure FTLD [141]. As *UBQLN2* is located on the X-chromosome, familial cases involving *UBQLN2* variants are characterized by an absence of observed male-to-male transmission in families. Although disease-linked variants in *UBQLN2* are rare in all populations studied to date [142–144], the fact that these variants impair protein degradation as well as the presence of ubiquilin 2 pathology in ALS and ALS-FTLD cases regardless of *UBQLN2* variant status [140] further highlights the importance of proper protein degradation in the ALS-FTLD disease spectrum.

Along the same lines, pathogenic variants in *SQSTM1* (sequestosome 1) have also been linked to familial and sporadic ALS cases [145]. *SQSTM1* encodes the ubiquitin-binding p62 protein, which is involved in targeting other proteins for autophagy. Additional studies have validated this genetic finding in ALS [146–150], and expanded it to link pathogenic variants in *SQSTM1* to ALS-FTLD [151] and pure FTLD [146] cases as well. Ubiquitin-binding p62 protein is found in ubiquitin-positive inclusions in several forms of ALS, including familial ALS due to pathogenic variants in *SOD1* [145]. Variants in *SQSTM1* account for ~1-5% of familial ALS cases and ~2-4% of sporadic cases

[150]. Interestingly, *SQSTM1* is also implicated in Paget's disease of bone [152], which is part of the IBMPFD phenotype associated with pathogenic variants in *VCP*.

Another shared genetic risk factor between the ALS-FTLD spectrum and Paget's disease of bone is *OPTN* (optineurin), which encodes for the optineurin protein and is involved in autophagy [153–155]. Finally, rare pathogenic variants in several additional genes have been implicated in the ALS-FTLD spectrum, including *TBK1* (TANK binding kinase 1), a gene encoding a kinase that binds to adaptor proteins including optineurin [156–159], *CHCHD10* (coiled-coil-helix-coiled-coil-helix domain containing 10), a gene involved in mitochondrial function [160–164], and *CHMP2B* (charged multivesicular body protein 2b), a gene involved in the endosomal/lysosomal degradation pathway [165–169]. For further descriptions and details of the genetic risk factors associated with the ALS-FTLD spectrum of disorders please refer to many detailed reviews [93,98,116,129]. Recent publications highlight the oligogenic hypothesis for ALS-FTLD, which suggests that variants in multiple loci contribute to both disease risk and clinical presentation [80,170].

FTLD: MAPT and GRN

Just as *SOD1* variants have only been linked with ALS risk and have not been implicated in FTLD, several genes have been linked to FTLD risk without being associated with ALS risk. Early investigations of the genetic contributions to a subtype of FTLD known as FTLD with parkinsonism identified the 17q21 region on chromosome 17, which was subsequently fine-mapped to two distinct risk loci. The first FTLD-causative pathogenic variants were identified in *MAPT* (microtubule associated protein tau), encoding for the tau protein, in 1998 [171–174]. As of 2017, 44 pathogenic variants

in *MAPT* have been identified (for an updated count, see the AD & Frontotemporal Dementia Mutation Database at http://www.molgen.ua.ac.be/FTDMutations). As mentioned previously, tau is one of the primary proteins present in pathogenic inclusions in brains of FTLD patients, with 45% of FTLD cases positive for tau protein inclusions [79]. While *MAPT* variants are not risk factors for ALS, the association of *MAPT* variants with PD risk and the presence of pathological aggregates of tau protein in AD again underscore the shared genetic etiology and pathological processes underlying several neurodegenerative diseases [2].

Surprisingly, after the mapping of *MAPT*, some families with FTLD with parkinsonism linked to chromosome 17 did not have pathogenic variants in *MAPT* and lacked tau inclusions upon pathological examination, suggesting another disease-causing locus existed outside of *MAPT*, yet within the same chromosomal region. Subsequent family studies identified pathogenic variants in *GRN* (granulin precursor, encoding progranulin protein), within the 17q21 region and just 1.7 mega bases from *MAPT*, that segregate with affected individuals [175,176]. After repeat expansions in *C9ORF72*, *GRN* variants are one of the most common genetic causes of both sporadic and familial FTLD [98], accounting for 5-10% of all FTLD cases and 10-22% of familial FTLD [177]. While the function of the progranulin protein is an area of active research, the majority of the 70 pathogenic variants within *GRN* that contribute to FTLD risk are thought to do so through loss of progranulin function due to haploinsufficiency [178]. *GRN* pathogenic variants have also been linked with AD (both phenotypic presentation and pathologically confirmed cases) in rare instances [179], possibly driven by presence of AD risk factor *APOE*-ε4 [180].

Historically, the clinical and pathological heterogeneity of FTLD has made it difficult to use GWAS to identify common genetic risk factors for this disorder. However, a recent GWAS focusing on pathologically confirmed FTLD cases with TDP-43 pathology identified a SNP tagging the *TMEM106B* (transmembrane protein 106B) gene associated with FTLD risk [181]. Subsequent studies have not only shown that SNPs in *TMEM106B* modulate risk for FTLD-TDP, but also regulate the penetrance of pathogenic *GRN* variants [182] and can decrease age of disease onset by a mean of 13 years in *GRN* pathogenic variant carriers [182,183]. Furthermore, *TMEM106B* variants can also modify disease risk in *C9ORF72* repeat expansion carriers – the minor allele is protective for FTD but not ALS [184], though interestingly the variant that increases FTD risk is associated with later age of onset and death in *C9ORF72* repeat expansion carriers [185]. While *TMEM106B* by itself has not been implicated in risk for ALS, variants within this gene have been associated with cognitive impairment in ALS cases [186]. Additional GWAS have identified other genetic variants that contribute to FTD; given the GWAS findings from PD and AD, the finding of FTD association with a SNP localized in the HLA region is of particular relevance [187].

Genetic contributions to the ALS-FTLD spectrum converge on a critical role of RNA and protein homeostasis in disease pathogenesis

The genetic variants associated with ALS and FTLD risk have contributed greatly to our understanding of the pathobiology underlying this spectrum of diseases, highlighting the importance of several biological processes in the etiology of ALS and FTLD. Chief among these are, at the broadest level, RNA homeostasis and protein homeostasis. TDP-43 and FUS are both involved in numerous processes related to RNA

homeostasis, including association with transcriptional machinery, regulation of splicing, transcriptional inhibition, and many other processes [129]. Combined, these two proteins regulate hundreds of downstream RNA targets, implicating genetic variation within *TARDBP* and *FUS* in a multitude of cellular processes that may be important in ALS-FTLD disease etiology. *C9ORF72* is also hypothesized to contribute to disease pathogenesis through RNA dysregulation, among other processes, as previously mentioned.

Many of the other genes involved in the ALS-FTLD spectrum are involved in protein homeostasis, particularly protein degradation and lysosomal/endosomal function. *VCP* is critical in protein quality control through the protein's involvement in the ubiquitin-proteasome complex, which targets proteins for degradation or recycling. *UBQLN2* is also involved in ubiquitin-mediated protein degradation, particularly in the transport of ubiquitin-tagged proteins targeted for destruction to the proteasome, as is the p62 protein encoded by *SQSTM1*. Both p62 and optineurin (encoded by *OPTN*) are involved in physically binding proteins targeted for autophagy to the autophagosome, and the TBK1 protein binds to optineurin and other adaptor proteins. CHMP2B is involved in sorting and transporting ubiquitinated targets to autophagosomes [129]. Genetic variation within these genes that are associated with risk for the ALS-FTLD disease spectrum highlights the critical role of dysregulated protein homeostasis — particularly within the processes of protein degradation and recycling — in disease pathogenesis. Targeting deficits in protein degradation is one area of therapeutic development for ALS-FTD, though further work is required to elucidate the molecular

mechanisms by which pathogenic variants in each of these genes specifically mechanistically contribute to disease.

Alzheimer's disease

AD is the most common cause of dementia, accounting for approximately 60-80% of dementia cases [188]. AD causes progressive cognitive decline, eventually resulting in the inability to complete everyday tasks, and ultimately death. It is the sixth leading cause of death in the United States and affects more than 5 million Americans, with prevalence expected to nearly triple to 13.8 million by 2050 [188]. AD is characterized by the accumulation of two pathological proteins: amyloid beta, which aggregates into amyloid plaques outside neurons, and tau, which collects into neurofibrillary tangles of hyper-phosphorylated protein within neurons.

Age is the greatest risk factor for AD [188], followed by family history, which suggests that disease risk is heritable. AD cases are classified into two main categories based on age of onset: early-onset AD (EOAD), occurring before age 65, and late-onset AD (LOAD), occurring at or after age 65. LOAD represents the vast majority of AD cases, with estimates ranging from 90% to 98% of AD cases [189,190]. In the early 1990s, epidemiological studies found an increased relative risk of 3.5 for AD in individuals with a family history of AD in first-degree relatives [191]. Early-onset cases in particular tend to be familial and highly heritable, with approximately 60% of cases occurring across at least two generations of a family [192]. The pathogenic variants and genes linked to early-onset familial AD (EOFAD) were the first genetic risk factors discovered for AD. Linkage analysis and fine-mapping studies in EOFAD families initially identified pathogenic variants in three genes linked to amyloid synthesis and

processing that segregated with affected individuals. These findings heavily influenced the focus of AD research over the past several decades, particularly with respect to the "amyloid cascade" hypothesis [193]. However, recent failures of many large-scale AD clinical trials with amyloid-targeted therapies are prompting a re-evaluation of the importance of amyloid beta in AD.

EOFAD: APP

Following studies that implicated chromosome 21 in AD, a point mutation in *APP* (amyloid precursor protein) was found to cause EOFAD in a British family [194,195]. This pathogenic variant was termed the "London mutation," and was followed up by additional studies that confirmed pathogenic variants in *APP* cause EOFAD [196–198]. *APP* is located on chromosome 21, and encodes the amyloid precursor protein. Proteolytic cleavage of APP produces a variety of peptides, including amyloid beta, which aggregates into amyloid plaques to form one of the two main pathological hallmarks of AD. Functional analyses of these disease-causing variants identified a shift in APP cleavage that favored the production of a more pathogenic form of amyloid, $A\beta_{42}$ [199]. To date, 51 pathogenic variants in *APP* have been identified (for an updated count, see the AD Mutation Database at http://www.molgen.vib-ua.be/ADMutations). Variants in *APP* account for approximately 10-15% of EOFAD [192], and the typical age of onset for pathogenic variant carriers is in their 40s-50s, with a few individuals having disease onset in their 60s [192]. While the majority of *APP* variants lead to early onset disease, rare *APP* variants have also been associated with LOAD [200].

Finding variants in *APP* associated with AD helped explain earlier findings from the mid-1980s showing that individuals with Down's syndrome (DS) developed amyloid

plaques and tau tangles approximately 20-30 years earlier than non-DS populations, and that dementia was clinically diagnosed in DS patients 20-30% more frequently than the general population [201]. DS is caused by the presence of a full or partial extra copy of chromosome 21, which includes the *APP* gene. Mouse models of DS have provided evidence that overexpressing *APP* contributes to the early pathological and clinical features of AD in DS [202].

More recently, whole-genome sequencing of 1,795 Icelandic individuals through the deCODE project identified a novel variant in *APP* (A673T) that protects against AD risk [203]. This variant has a large effect size (OR = 5.29 for control group vs. AD), and was shown to protect against age-related cognitive decline in non-diseased individuals, possibly by decreasing amyloid beta production [203]. As might be expected with such a large effect size, A673T occurs at low frequency in the Icelandic population (0.62% in controls, 0.13% in AD patients). This variant has not been observed in several Asian or non-Nordic Caucasian cohorts studied to date [204,205].

EOAD: PSEN1 and PSEN2

A few years after pathogenic variants in *APP* were linked to AD, two other components of the cellular amyloid processing machinery were implicated in AD: *PSEN1* (presenilin 1) and *PSEN2* (presenilin 2). The proteins encoded by these genes make up part of the gamma secretase enzyme complex, which cleaves APP to generate amyloid beta [206]. Pathogenic variants in these genes were initially identified in 1995 [207,208] and to date, 219 *PSEN1* and 16 *PSEN2* pathogenic variants have been identified [209,210]. Variants in *PSEN1* account for 30-70% of EOFAD [192], and lead to very severe clinical presentations of disease with high penetrance and an extremely early age of disease

onset, as early as 25 years [211]. While there is variability in disease severity, progression, and age of onset in *PSEN1* pathogenic variant carriers [211], on average these patients develop symptoms around age 40, which is over 8 years earlier than *APP* variant carriers, and over 14 years earlier than *PSEN2* variant carriers [209]. Variants in *PSEN2* account for less than 5% of EOFAD cases and display large variability in age of onset (typically 40-75 years), clinical presentation, and disease penetrance, with some individuals surviving past age 80 without developing symptoms [192]. Analysis of 128 individuals in the Dominantly Inherited Alzheimer Network cohort who all carried a pathogenic variant in either *PSEN1*, *PSEN2,* or *APP* found increased hippocampal atrophy, decreased cerebral glucose metabolism, increased amyloid deposition, and increased tau in the cerebrospinal fluid occurring 10-15 years prior to expected symptom onset in all variant carriers [212]. To date, variation in *PSEN1*, *PSEN2*, and *APP* are considered the major genetic risk factors for EOFAD, but additional genetic contributions to EOFAD risk are likely to exist since not all EOAD patients carry a pathogenic variant in one of these three genes. Furthermore, not all EOAD cases are familial, and the genetic contribution to these non-familial EOAD cases is poorly understood.

LOAD: APOE-ε4

The vast majority (> 95%) of clinical Alzheimer's cases are diagnosed at the age of 65 or later and are considered late-onset. The first genetic risk factor to be identified for these cases was *APOE-ε4* (apolipoprotein E4), and even after the discovery of over 20 additional genetic variants associated with LOAD risk, *APOE-ε4* remains the most common genetic risk factor of large effect in LOAD cases.

Concurrent with the discovery of *APP* variants causing AD, several linkage studies in families with late-onset disease failed to establish a link between disease risk and chromosome 21, suggesting that additional genetic risk factors beyond *APP* were contributing to AD risk [213]. Further studies implicated chromosome 19, and in 1993 a small study found that AD patients had a much higher frequency of the *APOE-ε4* allele than controls (0.50 vs. 0.16) [214]. Within half a year, two additional studies demonstrated the association of *APOE-ε4* with AD [215,216], and with subsequent studies, it has become clear that *APOE-ε4* is the most common genetic risk factor for AD and exerts a strikingly strong effect size, particularly for such a common variant.

There are three alleles of the *APOE* gene: *ε2*, which is very rare and is suggested to be protective against AD; *ε3*, the most common allele, considered to be 'neutral' (neither protective nor conferring risk); and *ε4*. Carrying the *APOE-ε4* allele increases the risk of developing AD in a dose-dependent manner, with one copy of *ε4* increasing the odds of developing AD by 2.7-3.2 fold and two copies of *ε4* increasing the odds of disease 12.5-14.9 fold [217]. The *APOE-ε4* allele also decreases the average age of onset, with one copy of *ε4* decreasing the age of onset by ~9 years and two copies of *ε4* decreasing the age of onset by an additional ~7 years [215,217]. *APOE-ε4* also has similar effects on average survival, as carrying the *APOE-ε4* allele is associated with shorter survival by 6.1 – 6.8 years [215]. While frequency of the *APOE-ε4* allele varies with both ancestry and latitude [218], approximately 23% of the US population carries at least one copy of the *APOE-ε4* allele [219]. *APOE-ε4* allele carriers are highly enriched in AD populations, with ~40-60% of AD patients carrying at least one *APOE-ε4* allele [220]. The *APOE-ε4* allele effect size also appears to vary by sex.

While women, regardless of genotype, are more likely to develop AD [217], the *APOE-ε4* allele also appears to confer greater AD risk in women [219]. Functionally, *APOE-ε4* is hypothesized to contribute to AD risk through both amyloid beta-dependent and amyloid beta-independent mechanisms [221,222]. In particular, a single amino acid difference between the *APOE-ε3* protein and the *APOE-ε4* protein allows for greater interaction between the N-terminal and C-terminal domains of the *APOE-ε4* protein, making this protein more susceptible to proteolytic cleavage, which in turn generates neurotoxic fragments that lead to cytoskeletal alterations and mitochondrial dysfunction [223]. Beyond increasing the risk of LOAD, *APOE-ε4* has been shown to influence age of onset of both *PSEN1* [224] and *PSEN2* variant carries [225]. Without a doubt, the *APOE-ε4* allele is one of the strongest, most well established genetic risk factors for LOAD.

LOAD: TREM2

While recent GWAS have identified a number of common variants associated with AD risk (as described below), these variants typically have relatively small effect sizes and modest contributions to disease risk relative to *APOE-ε4*. In contrast to the identification of common variants, the deCODE project (the same study that identified the protective *APP* variant described above) and another group concurrently identified rare variants in *TREM2* (triggering receptor expressed on myeloid cells 2), in particular the R47H variant, as risk factors for AD [226,227]. While somewhat rare (e.g. frequency of R47H = 0.26% [228]), these variants tend to be more common than the rare pathogenic variants in *APP*, *PSEN1,* or *PSEN2* linked to EOFAD [229]. *TREM2* encodes for a receptor expressed on myeloid cells, and is expressed in microglial cells in the brain. Combined

with the relatively large effect size of R47H (odds ratio of 4.5 (95% confidence interval 1.7-11.9) in AD cases vs. controls [230]), this discovery has ignited a flurry of studies focusing on the role of microglial activation and neuroinflammation in AD pathophysiology. Interestingly, rare homozygous recessive pathogenic variants in *TREM2* have also been linked to FTD and Nasu-Hakola disease, a rare degenerative bone condition [231,232].

LOAD: Additional common variants

The first GWAS for AD was conducted in a relatively small cohort of 1,808 LOAD cases and 2,062 older adult controls without dementia in the UK, and was only able to identify SNPs within *APOE* as significantly associated with AD risk [233]. Further GWAS using larger study cohorts and SNP panels with greater genomic coverage have identified and replicated a number of additional variants associated with AD risk, and are summarized elsewhere [234]. A recent meta-analysis of GWAS that included 74,046 individuals from 4 prior studies as well as analysis of new cases and controls from 11 different countries identified several new loci associated with AD, and replicated ten SNPs associated with AD risk in prior GWAS near the following genes: *ABCA7*, *APOE*, *BIN1*, *CD33*, *CLU*, *CR1*, *CD2AP*, *EPHA1*, *MS4A6A*, *PICALM* [235]. The most statistically significant of the novel findings was within the *HLA-DRB5/DRB1* region, followed by SNPs in close proximity to *SORL1*, *PTK2B*, *SLC24A4*, *ZCWPW1*, *CELF1*, *NME8*, *FERMT2*, *CASS4*, *INPP5D,* and *MEF2C*.

The loci and proximal genes identified by AD meta-analysis and other studies can generally be clustered into several functionally related groups reflecting pathways and processes that contribute to AD risk, such as immune response, endocytosis,

cholesterol metabolism, and ubiquitination. Furthermore, genes involved in the immune system and cholesterol metabolism are significantly overrepresented in GWAS for AD [236]. Of particular relevance to this chapter, the categories of immune response, endocytosis, and ubiquitination are shared amongst other neurodegenerative disorders. For example, while fine-mapping the HLA risk association has been difficult, recent studies have identified an association between several HLA haplotypes and AD risk [237], with numerous on-going fine-mapping efforts attempting to identify the specific HLA genes and/or haplotypes conferring risk of AD. It is also important to recognize that beyond genetic association studies, *in vitro* and *in vivo* experimental studies have confirmed the importance of the immune system in AD [238] and this continues to be a growing area of research in AD.

Because of the polygenic nature of this complex disease, some groups have begun using information from multiple disease-associated variants, with and without the inclusion of *APOE-ε4*, to make predictions of an individual's disease risk [239–244]. Very recently, Desikan and colleagues have identified a polygenic hazard score that predicts an individual's age-specific LOAD risk based on a combination of common genetic loci (including *APOE-ε4*) and population-level age-associated AD risk [245].

The paradox of AD genetics: Two pathological proteins, a variety of genetic risk factors, and no disease-modifying treatments

While the neuropathological diagnosis of AD requires the presence of two pathogenic protein aggregates (amyloid beta plaques and tau neurofibrillary tangles), there is debate in the AD field about whether the disease process is driven primarily by one of these two proteins, or even by *APOE-ε4* itself, either independently or in conjunction

with amyloid and/or tau. Therapeutic development thus far has focused primarily on reducing amyloid accumulation in the brain. These approaches have proved largely ineffective, as the majority of drugs targeting amyloid have failed in the clinical trial stage (though a few amyloid-focused trials with promising interim results are currently underway) [246]. The discovery of genetic variants in the amyloid processing pathway (in *APP*, *PSEN1* and *PSEN2*) with extremely high penetrance is likely an unfortunate reason for this singular focus on amyloid in the AD field. These pathogenic variants were an extremely exciting and early discovery in the AD field and led to the "amyloid cascade" hypothesis, which posits that aberrant APP processing leads to pathogenic amyloid beta accumulation which, in turn, causes neuronal loss and cognitive deficits [193,247].

The amyloid cascade hypothesis has led to animal models and experiments that have informed our understanding of how dysregulation in the amyloid processing pathway can lead to cognitive deficits, particularly in the subset of patients that carry pathogenic variants in genes in this pathway. Although this line of biological inquiry is highly relevant for individuals harboring pathogenic variants in *APP*, *PSEN1* and *PSEN2,* these are overall a very rare cause of AD. In non-genetic and late-onset forms of AD, it is becoming increasingly clear that the majority of AD cases are not driven solely — or perhaps even primarily — by amyloid aggregation. For example, in the Clifford Jack model of disease pathogenesis and biomarker status, amyloid aggregation is the earliest step in AD pathogenesis, and plaque formation is necessary but not sufficient for clinical AD. Amyloid beta deposition is followed by increases in measures of pathogenic tau, which in turn is followed by structural and metabolic imaging changes

and, eventually, by cognitive decline [248]. Proponents of the amyloid cascade hypothesis argue that future trials focusing on disease prevention, rather than symptomatic treatment or reversal, may show more promising effects; the results of these trials, many of which are currently underway, will inform our understanding of the validity of the "amyloid cascade" hypothesis in AD etiology. Based on the Jack model of AD progression, however, preventing AD through amyloid-directed therapies may require individuals to be treated up to 20 years prior to symptom onset. While identifying genetically at-risk individuals is one way to do this, other strategies to modify disease trajectory at different stages of AD — e.g., prior to tau deposition or neuronal loss — may prove alternative or complementary strategies to modifying AD. Therapeutic strategies targeting tau or *APOE-ε4* are currently in the preclinical and/or early clinical stages; the results of these trials over the next five to ten years will provide great insight into the pathophysiology of AD, especially in assessing the contributions of amyloid, tau, and *APOE-ε4* to disease. Work also suggests that preventing aberrant neuronal function (e.g., epileptiform activity [249,250]), neuronal pruning [251], and augmenting synaptic plasticity [252–254] may promote maintenance of cognitive function in the face of burgeoning disease pathology.

Conclusions

There is considerable overlap between genetic risk factors and disease presentation among the neurodegenerative diseases described in this chapter. Several of these diseases can be caused by autosomal dominant pathogenic variants in a single gene (such as *HTT* in HD, *SCNA* in PD, *C9ORF72* and *TARDBP* in ALS-FTD, *MAPT* in FTLD, and *APP* in AD). These pathogenic variants lead to deleterious accumulation of

a single protein (usually a neuropathological hallmark of that disease), though the prevalence of these mutations ranges from all patients (as in *HTT* pathogenic variants in HD) to only very rare cases of a given disease (like *APP* pathogenic variants in AD). While these autosomal dominant, disease-causing variants can provide insight into the etiology of neurodegenerative diseases, we have yet to develop effective therapies targeting any of these genetic causes of disease, and in some cases, the concentrated focus on a single disease-causing gene (at the expense of focus on the plethora of genetic regions that can contribute to disease risk) may even be leading us astray, as may be the case with the focus on amyloid beta in AD.

More often, neurodegenerative diseases may arise through a combination of subtle effects from many genes that have been linked to disease risk. While the exact genes that contribute to disease risk vary between diseases, many genes involved in RNA homeostasis, protein homeostasis, mitochondrial function, and the immune system have been implicated in one or more of the neurodegenerative diseases described in this chapter. As demonstrated on numerous occasions, genetic variation within a single region can also confer risk for multiple distinct neurodegenerative diseases, such as *MAPT* (FTD and FTLD, PD) and the HLA region (PD, FTD, AD). This so-called "pleiotropy" can be detected statistically [2,255,256] and may have implications for "pan neurodegenerative" strategies for disease treatment.

Beyond the overlap of individual risk loci between numerous neurodegenerative diseases, diverse loci are also associated with shared pathways and processes between different neurodegenerative diseases, suggesting that these discrete disorders may have common mechanistic underpinnings. Chief among these are aggregation of a

pathological protein (such as α-synuclein, TDP-43, tau, or amyloid beta) and impairments in protein degradation or autophagy, which can compound the deleterious effects of pathogenic protein aggregation. Both endocytosis and ubiquitination are involved in cellular proteostasis, which is hypothesized to be dysregulated in several neurodegenerative diseases. Growing evidence has shown that pathogenic proteins such as tau (in AD and FTLD), TDP-43 (in ALS and FTLD) and alpha-synuclein (in PD) can be transmitted from cell to cell within the brain and even outside the brain. This process clearly necessitates the internalization of these proteins by the receiving cell, which most likely occurs through endocytosis. Once a pathogenic protein has entered a cell, deficits in proteolysis linked to lysosomal or ubiquitination defects likely prevent its destruction, and allow the protein's pathogenic effects to proliferate and spread. The involvement of ubiquitination in the neurodegenerative diseases described within this chapter is described in detail in Atkin & Paulson's review [257].

Altered immune system function has also been implicated in several neurodegenerative disorders, and may play an important role in the brain's response to pathological protein aggregates. The immune system has been a growing area of interest in AD research in particular, with recent studies finding genetic polymorphisms linked to autoimmune disorders are often also associated with AD risk [255], and validating previous studies identifying variation in genes such as *CD33* (a myeloid cell surface receptor), *TREM2* (described in detail above, receptor on myeloid cells), *CR1* (complement receptor 1), and the *HLA* (human leukocyte antigen) region as risk factors for AD. As mentioned previously, the identification of variants within the HLA region associated with PD and FTD risk in addition to AD highlights the importance of this

association across the spectrum of neurodegenerative diseases, though whether all diseases are linked to a single risk locus or different risk loci within this region remains to be elucidated.

While translational research aimed at discovering therapies for neurodegenerative diseases has historically focused on genetic risk factors for these diseases, one area of growing interest is the focus on genetic factors that contribute to healthy cognitive aging, with the hope that genetic variation linked to healthy aging may provide alternative therapeutic targets [258–265]. While a description of these factors is beyond the scope of this chapter, readers may find some useful information in reviews of healthy cognitive aging [266] or more general reviews of healthy aging and longevity [267,268].

As biomedical research provides solutions for common health problems such as heart disease and cancer, neurodegenerative diseases will become increasingly more prevalent due to our aging population. The enormous burden this will create on our society makes developing therapies for these devastating diseases increasingly important. Of course, it is important to keep in mind that genetic variation is only one of many factors that contribute to overall disease risk. However, research has shown that the likelihood of success for a therapeutic approach is much higher when supported by evidence of a genetic factor contributing to disease risk in humans [269]. Further exploration of the mechanistic contributions of established genetic risk factors to the pathobiology of neurodegenerative diseases, in addition to the identification of additional novel genetic risk factors for these disorders, will continue to provide invaluable insight into how genetics influences the process of brain aging.

References

1. Desikan RS, Schork AJ, Wang Y, Witoelar A, Sharma M, McEvoy LK, et al. Genetic overlap between Alzheimer's disease and Parkinson's disease at the MAPT locus. Mol Psychiatry. 2015; 1–8. doi:10.1038/mp.2015.6

2. Ferrari R, Wang Y, Vandrovcova J, Guelfi S, Witeolar A, Karch CM, et al. Genetic architecture of sporadic frontotemporal dementia and overlap with Alzheimer's and Parkinson's diseases. J Neurol Neurosurg Psychiatry. 2016; jnnp-2016-314411. doi:10.1136/jnnp-2016-314411

3. Yokoyama JS, Karch CM, Fan CC, Bonham LW, Kouri N, Ross OA, et al. Shared genetic risk between corticobasal degeneration, progressive supranuclear palsy, and frontotemporal dementia. Acta Neuropathol. 2017; 1–13. doi:10.1007/s00401-017-1693-y

4. Warby SC, Graham RK, Hayden MR. Huntington Disease. GeneReviews™ [Internet]. 2014;58: 308–310. doi:http://www.ncbi.nlm.nih.gov/books/NBK1305/

5. Walker FO. Huntington's disease. Lancet. 2007. pp. 218–228. doi:10.1016/S0140-6736(07)60111-1

6. Gusella JF, Wexler NS, Conneally PM, Naylor SL, Anderson M a, Tanzi RE, et al. A polymorphic DNA marker genetically linked to Huntington's disease. Nature. 1983;306: 234–238. doi:10.1017/CBO9781107415324.004

7. MacDonald ME, Ambrose CM, Duyao MP, Myers RH, Lin C, Srinidhi L, et al. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. Cell. 1993;72: 971–983. doi:10.1016/0092-8674(93)90585-E

8. Kremer B, Goldberg P, Andrew SE, Theilmann J, Telenius H, Zeisler J, et al. A Worldwide Study of the Huntington's Disease Mutation: The Sensitivity and Specificity of Measuring CAG Repeats. N Engl J Med. 1994;330: 1401–1406. doi:10.1056/nejm199405193302001

9. Semaka A, Creighton S, Warby S, Hayden MA. Predictive testing for Huntington disease: Interpretation and significance of intermediate alleles. Clinical Genetics. 2006. pp. 283–294. doi:10.1111/j.1399-0004.2006.00668.x

10. Semaka A, Kay C, Doty C, Collins JA, Bijlsma EK, Richards F, et al. CAG size-specific risk estimates for intermediate allele repeat instability in Huntington disease. J Med Genet. 2013;50: 696–703. doi:10.1136/jmedgenet-2013-101796

11. Semaka A, Balneaves LG, Hayden MR. "Grasping the grey": Patient understanding and interpretation of an intermediate allele predictive test result for huntington disease. J Genet Couns. 2013;22: 200–217. doi:10.1007/s10897-012-9533-7

12. Myers RH. Huntington's disease genetics. NeuroRx. 2004;1: 255–62. doi:10.1602/neurorx.1.2.255

13. Goldberg YP, Kremer B, Andrew SE, Theilmann J, Graham RK, Squitieri F, et al. Molecular analysis of new mutations for Huntington's disease: intermediate alleles and sex of origin effects. Nat Genet. 1993;5: 174–9. doi:10.1038/ng1093-174

14. Zühlke C, Rless O, Bockel B, Lange H, Thies U. Mitotic stability and meiotic variability of the (CAG)n repeat in the huntington disease gene. Hum Mol Genet. 1993;2: 2063–2067. doi:10.1093/hmg/2.12.2063

15. Brinkman RR, Mezei MM, Theilmann J, Almqvist E, Hayden MR. The likelihood of being affected with Huntington disease by a particular age, for a specific CAG size. Am J Hum Genet. 1997;60: 1202–10. doi:.

16. Rubinsztein DC, Leggo J, Coles R, Almqvist E, Biancalana V, Cassiman JJ, et al. Phenotypic characterization of individuals with 30-40 CAG repeats in the Huntington disease (HD) gene reveals HD cases with 36 repeats and apparently normal elderly individuals with 36-39 repeats. Am J Hum Genet. 1996;59: 16–22.

17. McNeil SM, Novelletto a, Srinidhi J, Barnes G, Kornbluth I, Altherr MR, et al. Reduced penetrance of the Huntington's disease mutation. Hum Mol Genet. 1997;6: 775–779. doi:10.1093/hmg/6.5.775

18. Langbehn DR, Brinkman RR, Falush D, Paulsen JS, Hayden MR. A new model for prediction of the age of onset and penetrance for Huntington's disease based on CAG length. Clin Genet. 2004;65: 267–277. doi:10.1111/j.1399-0004.2004.00241.x

19. Duyao M, Ambrose C, Myers R, Novelletto A, Persichetti F, Frontali M, et al. Trinucleotide repeat length instability and age of onset in Huntington's disease. Nat Genet. 1993;4: 387–392. doi:10.1038/ng0893-387

20. Andrew SE, Goldberg YP, Kremer B, Telenius H, Theilmann J, Adam S, et al. The relationship between trinucleotide (CAG) repeat length and clinical features of Huntington's disease. Nat Genet. 1993;4: 398–403. doi:10.1038/ng0893-398

21. Nance MA, Myers RH. Juvenile onset Huntington's disease--clinical and research perspectives. Ment Retard Dev Disabil Res Rev. 2001;7: 153–7. doi:10.1002/mrdd.1022

22. Bates GP. History of genetic disease: The molecular genetics of Huntington disease – a history. Nat Rev Genet. 2005;6: 766–773. doi:10.1038/nrg1686

23. Gusella JF, MacDonald ME, Lee J-M. Genetic modifiers of Huntington's disease. Mov Disord. 2014;29: 1359–65. doi:10.1002/mds.26001

24. Lee JM, Wheeler VC, Chao MJ, Vonsattel JPG, Pinto RM, Lucente D, et al. Identification of Genetic Factors that Modify Clinical Onset of Huntington's Disease. Cell. 2015;162: 516–526. doi:10.1016/j.cell.2015.07.003

25. Finkbeiner S. Huntington's disease. Cold Spring Harb Perspect Biol. 2011;3: 1–24. doi:10.1101/cshperspect.a007476

26. Scherzinger E, Lurz R, Turmaine M, Mangiarini L, Hollenbach B, Hasenbank R, et al. Huntingtin-encoded polyglutamine expansions form amyloid-like protein aggregates in vitro and in vivo. Cell. 1997;90: 549–558. doi:10.1016/S0092-8674(00)80514-0

27. Labbadia J, Morimoto RI. Huntington's disease: Underlying molecular mechanisms and emerging concepts. Trends in Biochemical Sciences. 2013. pp. 378–385. doi:10.1016/j.tibs.2013.05.003

28. DeJesus-Hernandez M, Mackenzie IR, Boeve BF, Boxer AL, Baker M, Rutherford NJ, et al. Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. Neuron. Department of Neuroscience, Mayo Clinic Florida, Jacksonville, FL 32224, USA.; 2011;72: 245–256. Available: http://www.sciencedirect.com/science/article/pii/S0896627311008282

29. Lill CM. Genetics of Parkinson's disease. Molecular and Cellular Probes. 2016. pp. 386–396. doi:10.1016/j.mcp.2016.11.001

30. Gelb DJ, Oliver E, Gilman S. Diagnostic criteria for Parkinson disease. Arch Neurol. 1999;56: 33–9. doi:10.1001/archneur.56.1.33

31. Trinh J, Farrer M. Advances in the genetics of Parkinson disease. Nat Rev Neurol. 2013;9: 445–54. doi:10.1038/nrneurol.2013.132

32. Klein C, Westenberger A, Hollingworth P, Harold D, Jones L, Owen MJ, et al. Genetics of Parkinson's Disease. Int J Geriatr Psychiatry. 2012;26: a008888. doi:10.1101/cshperspect.a008888

33. Polymeropoulos MH, Lavedan C, Leroy E, Ide SE, Dehejia A, Dutra A, et al. Mutation in the α-Synuclein Gene Identified in Families with Parkinson's Disease. Science (80-). 1997;276: 2045–2047. doi:10.1126/science.276.5321.2045

34. Spillantini MG, Schmidt ML, Lee VM, Trojanowski JQ, Jakes R, Goedert M. Alpha-synuclein in Lewy bodies. Nature. 1997;388: 839–840. doi:10.1038/42166

35. Krüger R, Kuhn W, Müller T, Woitalla D, Graeber M, Kösel S, et al. Ala30Pro mutation in the gene encoding alpha-synuclein in Parkinson's disease. Nat Genet. 1998;18: 106–108. doi:10.1038/ng0298-106

36. Zarranz JJ, Alegre J, Gómez-Esteban JC, Lezcano E, Ros R, Ampuero I, et al. The new mutation, E46K, of alpha-synuclein causes Parkinson and Lewy body dementia. Ann Neurol. 2004;55: 164–73. doi:10.1002/ana.10795

37. Appel-Cresswell S, Vilarino-Guell C, Encarnacion M, Sherman H, Yu I, Shah B, et al. Alpha-synuclein p.H50Q, a novel pathogenic mutation for Parkinson's disease. Mov Disord. 2013;28: 811–813. doi:10.1002/mds.25421

38. Proukakis C, Dudzik CG, Brier T, MacKay DS, Cooper JM, Millhauser GL, et al. A novel α-synuclein missense mutation in Parkinson disease. Neurology. 2013. pp. 1062–1064. doi:10.1212/WNL.0b013e31828727ba

39. Lesage S, Anheim M, Letournel F, Bousset L, Honoré A, Rozas N, et al. G51D ??-synuclein mutation causes a novel Parkinsonian-pyramidal syndrome. Ann Neurol. 2013;73: 459–471. doi:10.1002/ana.23894

40. Kiely AP, Asi YT, Kara E, Limousin P, Ling H, Lewis P, et al. A-synucleinopathy associated with G51D SNCA mutation: A link between Parkinson's disease and multiple system atrophy? Acta Neuropathol. 2013;125: 753–769. doi:10.1007/s00401-013-1096-7

41. Chartier-Harlin MC, Kachergus J, Roumier C, Mouroux V, Douay X, Lincoln S, et al. α-synuclein locus duplication as a cause of familial Parkinson's disease. Lancet. 2004;364: 1167–1169. doi:10.1016/S0140-6736(04)17103-1

42. Ibáñez P, Bonnet a-M, Débarges B, Lohmann E, Tison F, Pollak P, et al. Causal relation between alpha-synuclein gene duplication and familial Parkinson's disease. Lancet. 2004;364: 1169–1171. doi:10.1016/S0140-6736(04)17104-3

43. Singleton a B, Farrer M, Johnson J, Singleton a, Hague S, Kachergus J, et al. alpha-Synuclein locus triplication causes Parkinson's disease. Science. 2003;302: 841. doi:10.1126/science.1090278

44. Lill C, Mashychev A, Hartmann C, Lohmann K, Marras C, Lang A, et al. Launching the movement disorders society genetic mutation database (MDSGene). Mov Disord. 2016;31: 607–9. doi:10.1002/mds.26651

45. Nalls M a, Pankratz N, Lill CM, Do CB, Hernandez DG, Saad M, et al. Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. Nat Genet. 2014;56: 1–7. doi:10.1038/ng.3043

46. Lill CM, Roehr JT, McQueen MB, Kavvoura FK, Bagade S, Schjeide BMM, et al. Comprehensive research synopsis and systematic meta-analyses in Parkinson's disease genetics: The PDgene database. PLoS Genet. 2012;8. doi:10.1371/journal.pgen.1002548

47. Nalls MA, Plagnol V, Hernandez DG, Sharma M, Sheerin UM, Saad M, et al. Imputation of sequence variants for identification of genetic risks for Parkinson's disease: A meta-analysis of genome-wide association studies. Lancet. 2011;377: 641–649. doi:10.1016/S0140-6736(10)62345-8

48. Burré J, Sharma M, Tsetsenis T, Buchman V, Etherton MR, Südhof TC. Alpha-synuclein promotes SNARE-complex assembly in vivo and in vitro. Science. 2010;329: 1663–7. doi:10.1126/science.1195227

49. Zimprich A, Biskup S, Leitner P, Lichtner P, Farrer M, Lincoln S, et al. Mutations in LRRK2 cause autosomal-dominant parkinsonism with pleomorphic pathology. Neuron. Department of Neurodegenerative Disease, Hertie Institute for Clinical Brain Research, University of Tübingen, 72076 Tübingen, Germany.; 2004;44: 601–607.

50. Paisán-Ruíz C, Jain S, Evans EW, Gilks WP, Simón J, Van Der Brug M, et al. Cloning of the gene containing mutations that cause PARK8-linked Parkinson's disease. Neuron. 2004;44: 595–600. doi:10.1016/j.neuron.2004.10.023

51. Healy DG, Falchi M, O'Sullivan SS, Bonifati V, Durr A, Bressman S, et al. Phenotype, genotype, and worldwide genetic penetrance of LRRK2-associated Parkinson's disease: a case-control study. Lancet Neurol. 2008;7: 583–590. doi:10.1016/S1474-4422(08)70117-0

52. Ross OA, Soto-Ortolaza AI, Heckman MG, Aasly JO, Abahuni N, Annesi G, et al. Association of LRRK2 exonic variants with susceptibility to Parkinson's disease: A case-control study. Lancet Neurol. 2011;10: 898–908. doi:10.1016/S1474-4422(11)70175-2

53. Mullin S, Schapira A. The genetics of Parkinson ' s disease. Br Med Bull. 2015; 1–14. doi:10.1093/bmb/ldv022

54. Satake W, Nakabayashi Y, Mizuta I, Hirota Y, Ito C, Kubo M, et al. Genome-wide association study identifies common variants at four loci as genetic risk factors for Parkinson's disease. Nat Genet. 2009;41: 1303–1307. doi:ng.485 [pii] 10.1038/ng.485

55. Do C, Tung J, Dorfman E, Kiefer A, Drabant E, Francke U, et al. Web-Based Genome-Wide Association Study Identifies Two Novel Loci and a Substantial Genetic Component for Parkinson's Disease. PLoS Genet. 2011;7: e1002141.

56. Vilariño-Güell C, Wider C, Ross OA, Dachsel JC, Kachergus JM, Lincoln SJ, et al. VPS35 mutations in parkinson disease. Am J Hum Genet. 2011;89: 162–167. doi:10.1016/j.ajhg.2011.06.001

57. Zimprich A, Benet-Pagès A, Struhal W, Graf E, Eck SH, Offman MN, et al. A mutation in VPS35, encoding a subunit of the retromer complex, causes late-onset parkinson disease. Am J Hum Genet. 2011;89: 168–175. doi:10.1016/j.ajhg.2011.06.008

58. Bonifacino JS, Hurley JH. Retromer. Current Opinion in Cell Biology. 2008. pp. 427–436. doi:10.1016/j.ceb.2008.03.009

59. Small SA, Petsko GA. Retromer in Alzheimer disease, Parkinson disease and other neurological disorders. Nat Rev Neurosci. 2015;16: 126–132. doi:10.1038/nrn3896

60. Kilarski LL, Pearson JP, Newsway V, Majounie E, Knipe MDW, Misbahuddin A, et al. Systematic Review and UK-Based Study of PARK2 (parkin), PINK1, PARK7 (DJ-1) and LRRK2 in early-onset Parkinson's disease. Mov Disord. 2012;27: 1522–1529. doi:10.1002/mds.25132

61. Kitada T, Asakawa S, Hattori N, Matsumine H, Yamamura Y, Minoshima S, et al. Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. Nature. 1998;392: 605–608. doi:10.1038/33416

62. Valente EM, Abou-Sleiman PM, Caputo V, Muqit MMK, Harvey K, Gispert S, et al. Hereditary early-onset Parkinson's disease caused by mutations in PINK1. Science. 2004;304: 1158–60. doi:10.1126/science.1096284

63. Narendra DP, Youle RJ. Targeting Mitochondrial Dysfunction: Role for PINK1 and Parkin in Mitochondrial Quality Control. Antioxid Redox Signal. 2011;14: 1929–1938. doi:10.1089/ars.2010.3799

64. Bonifati V, Rizzu P, van Baren MJ, Schaap O, Breedveld GJ, Krieger E, et al. Mutations in the DJ-1 gene associated with autosomal recessive early-onset parkinsonism. Science. 2003;299: 256–259. doi:10.1126/science.1077209

65. Rizzu P, Hinkle DA, Zhukareva V, Bonifati V, Severijnen LA, Martinez D, et al. DJ-1 Colocalizes with Tau Inclusions: A Link between Parkinsonism and Dementia. Ann Neurol. 2004;55: 113–118. doi:10.1002/ana.10782

66. van der Merwe C, Jalali Sefid Dashti Z, Christoffels A, Loos B, Bardien S. Evidence for a common biological pathway linking three Parkinson's disease-causing genes: Parkin, PINK1 and DJ-1. European Journal of Neuroscience. 2015. pp. 1113–1125. doi:10.1111/ejn.12872

67. Edvardson S, Cinnamon Y, Ta-Shma A, Shaag A, Yim YI, Zenvirt S, et al. A deleterious mutation in DNAJC6 encoding the neuronal-specific clathrin-uncoating Co-chaperone auxilin, is associated with juvenile parkinsonism. PLoS One. 2012;7. doi:10.1371/journal.pone.0036458

68. Köroğlu Ç, Baysal L, Cetinkaya M, Karasoy H, Tolun A. DNAJC6 is responsible for juvenile parkinsonism with phenotypic variability. Park Relat Disord. 2013;19: 320–324. doi:10.1016/j.parkreldis.2012.11.006

69. Olgiati S, Quadri M, Fang M, Rood JPMA, Saute JA, Chien HF, et al. DNAJC6 Mutations Associated with Early-Onset Parkinson's Disease. Ann Neurol. 2016;79: 244–256. doi:10.1002/ana.24553

70. Hamza TH, Zabetian CP, Tenesa A, Laederach A, Montimurro J, Yearout D, et al. Common genetic variation in the HLA region is associated with late-onset sporadic Parkinson's disease. Nat Genet. 2010;42: 781–785. doi:10.1038/ng.642

71. Ahmed I, Tamouza R, Delord M, Krishnamoorthy R, Tzourio C, Mulot C, et al. Association between Parkinson's disease and the HLA-DRB1 locus. Mov Disord. 2012;27: 1104–1110. doi:10.1002/mds.25035

72. Gagliano SA, Pouget JG, Hardy J, Knight J, Barnes MR, Ryten M, et al. Genomics implicates adaptive and innate immunity in Alzheimer's and Parkinson's [Internet]. bioRxiv. 2016. doi:10.1101/059519

73. Healy DG, Abou-Sleiman PM, Lees AJ, Casas JP, Quinn N, Bhatia K, et al. Tau gene and Parkinson's disease: a case-control study and meta-analysis. J Neurol Neurosurg Psychiatry. 2004;75: 962–5. doi:10.1136/jnnp.2003.026203

74. Setó-Salvia N, Clarimón J, Pagonabarraga J, Pascual-Sedano B, Campolongo A, Combarros O, et al. Dementia risk in Parkinson disease: disentangling the role of MAPT haplotypes. Arch Neurol. 2011;68: 359–64. doi:10.1001/archneurol.2011.17

75. Ezquerra M, Pastor P, Gaig C, Vidal-Taboada JM, Cruchaga C, Muñoz E, et al. Different MAPT haplotypes are associated with Parkinson's disease and progressive supranuclear palsy. Neurobiol Aging. 2011;32. doi:10.1016/j.neurobiolaging.2009.09.011

76. Lin MK, Farrer MJ. Genetics and genomics of Parkinson's disease. Genome Med. 2014;6: 48. doi:10.1186/gm566

77. Al-Chalabi A, Hardiman O. The epidemiology of ALS: a conspiracy of genes, environment and time. Nat Rev Neurol. 2013;9: 617–28. doi:10.1038/nrneurol.2013.203

78. Taylor JP, Brown RH, Cleveland DW. Decoding ALS: from genes to mechanism. Nature. 2016;539: 197–206. doi:10.1038/nature20413

79. Ling S-C, Polymenidou M, Cleveland DW. Converging mechanisms in ALS and FTD: disrupted RNA and protein homeostasis. Neuron. 2013;79: 416–38. doi:10.1016/j.neuron.2013.07.033

80. Lattante S, Ciura S, Rouleau GA, Kabashi E. Defining the genetic connection linking amyotrophic lateral sclerosis (ALS) with frontotemporal dementia (FTD). Trends in Genetics. 2015. pp. 263–273. doi:10.1016/j.tig.2015.03.005

81. Sieben A, Van Langenhove T, Engelborghs S, Martin J-J, Boon P, Cras P, et al. The genetics and neuropathology of frontotemporal lobar degeneration. Acta Neuropathol. 2012;124: 353–72. doi:10.1007/s00401-012-1029-x

82. Rosness TA, Engedal K, Chemali Z. Frontotemporal Dementia: An Updated Clinicians Guide. J Geriatr Psychiatry Neurol. 2016;29: 271–280. doi:10.1177/0891988716654986

83. Rascovsky K, Hodges JR, Knopman D, Mendez MF, Kramer JH, Neuhaus J, et al. Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. Brain. 2011;134: 2456–2477. Available: http://www.brain.oxfordjournals.org/cgi/doi/10.1093/brain/awr179

84. Gorno-Tempini ML, Hillis AE, Weintraub S, Kertesz A, Mendez M, Cappa SF, et al. Classification of primary progressive aphasia and its variants. Neurology. 2011;76: 1006–1014.

85. Dickson DW, Kouri N, Murray ME, Josephs KA. Neuropathology of frontotemporal lobar degeneration-Tau (FTLD-Tau). Journal of Molecular Neuroscience. 2011. pp. 384–389. doi:10.1007/s12031-011-9589-0

86. Mackenzie IR, Neumann M, Baborie A, Sampathu D, Du Plessis D, Jaros E, et al. A harmonized classification system for FTLD-TDP pathology. Acta Neuropathol. 2011;122: 111–113.

87. Lomen-Hoerth C, Anderson T, Miller B. The overlap of amyotrophic lateral sclerosis and frontotemporal dementia. Neurology. 2002;59: 1077–1079. doi:10.1097/WCO.0b013e3283168d1d

88. Ringholz GM, Appel SH, Bradshaw M, Cooke NA, Mosnik DM, Schulz PE. Prevalence and patterns of cognitive impairment in sporadic ALS. Neurology. 2005;65: 586–590. doi:10.1212/01.wnl.0000172911.39167.b6

89. Wheaton MW, Salamone AR, Mosnik DM, McDonald RO, Appel SH, Schmolck HI, et al. Cognitive impairment in familial ALS. Neurology. 2007;69: 1411–1417. doi:10.1212/01.wnl.0000277422.11236.2c

90. Burrell JR, Kiernan MC, Vucic S, Hodges JR. Motor Neuron dysfunction in frontotemporal dementia. Brain. 2011;134: 2582–2594. doi:10.1093/brain/awr195

91. Neumann M, Sampathu DM, Kwong LK, Truax AC, Micsenyi MC, Chou TT, et al. Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Science. 2006;314: 130–3. doi:10.1126/science.1134108

92. Arai T, Hasegawa M, Akiyama H, Ikeda K, Nonaka T, Mori H, et al. TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Biochem Biophys Res Commun. 2006;351: 602–611. doi:10.1016/j.bbrc.2006.10.093

93. van Langenhove T, van der Zee J, van Broeckhoven C. The molecular basis of the frontotemporal lobar degeneration-amyotrophic lateral sclerosis spectrum. Ann Med. 2012;44: 817–828. doi:10.3109/07853890.2012.665471

94. James BD, Wilson RS, Boyle PA, Trojanowski JQ, Bennett DA, Schneider JA. TDP-43 stage, mixed pathologies, and clinical Alzheimer's-type dementia. Brain. 2016;139: 2983–2993. doi:10.1093/brain/aww224

95. Rosen DR, Siddique T, Patterson D, Figlewicz DA, Sapp P, Hentati A, et al. Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. Nature. 1993;362: 59–62. doi:10.1038/362059a0

96. Al-Chalabi A. GENE OVERVIEW OF KEY PUBLISHED STUDIES FOR ALS [Internet]. 2015. Available: http://alsod.iop.kcl.ac.uk/als/Overview/gene.aspx?gene_id=SOD1

97. Cudkowicz ME, McKenna-Yasek D, Sapp PE, Chin W, Geller B, Hayden DL, et al. Epidemiology of mutations in superoxide dismutase in amyotrophic lateral sclerosis. Ann Neurol. 1997;41: 210–221. doi:10.1002/ana.410410212

98. Bennion Callister J, Pickering-Brown SM. Pathogenesis/genetics of frontotemporal dementia and how it relates to ALS. Experimental Neurology. 2014. pp. 84–90. doi:10.1016/j.expneurol.2014.06.001

99. Renton AE, Chiò A, Traynor BJ. State of play in amyotrophic lateral sclerosis genetics. Nat Neurosci. 2014;17: 17–23. doi:10.1038/nn.3584

100. Mackenzie IRA, Bigio EH, Ince PG, Geser F, Neumann M, Cairns NJ, et al. Pathological TDP-43 distinguishes sporadic amyotrophic lateral sclerosis from amyotrophic lateral sclerosis with SOD1 mutations. Ann Neurol. 2007;61: 427–434. doi:10.1002/ana.21147

101. Chen S, Sayana P, Zhang X, Le W. Genetics of amyotrophic lateral sclerosis: an update. Mol Neurodegener. 2013;8: 28. doi:10.1186/1750-1326-8-28

102. Miller TM, Pestronk A, David W, Rothstein J, Simpson E, Appel SH, et al. An antisense oligonucleotide against SOD1 delivered intrathecally for patients with SOD1 familial amyotrophic lateral sclerosis: A phase 1, randomised, first-in-man study. Lancet Neurol. 2013;12: 435–442. doi:10.1016/S1474-4422(13)70061-9

103. Gitcho MA, Baloh RH, Chakraverty S, Mayo K, Norton JB, Levitch D, et al. TDP-43 A315T mutation in familial motor neuron disease. Ann Neurol. 2008;63: 535–538. doi:10.1002/ana.21344

104. Kabashi E, Valdmanis PN, Dion P, Spiegelman D, McConkey BJ, Vande Velde C, et al. TARDBP mutations in individuals with sporadic and familial amyotrophic lateral sclerosis. Nat Genet. 2008;40: 572–574. doi:10.1038/ng.132

105. Sreedharan J, Blair IP, Tripathi VB, Hu X, Vance C, Rogelj B, et al. TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. Science (80-). 2008;319: 1668–72. doi:10.1126/science.1154584

106. Van Deerlin VM, Leverenz JB, Bekris LM, Bird TD, Yuan W, Elman LB, et al. TARDBP mutations in amyotrophic lateral sclerosis with TDP-43 neuropathology: a genetic and histopathological analysis. Lancet Neurol. 2008;7: 409–416. doi:10.1016/S1474-4422(08)70071-1

107. Borroni B, Bonvicini C, Alberici A, Buratti E, Agosti C, Archetti S, et al. Mutation within TARDBP leads to frontotemporal dementia without motor neuron disease. Hum Mutat. Centre for Ageing Brain and Neurodegenerative Disorders, University of Brescia, Brescia, Italy. bborroni@inwind.it; 2009;30: E974-83.

108. Kovacs GG, Murrell JR, Horvath S, Haraszti L, Majtenyi K, Molnar MJ, et al. TARDBP variation associated with frontotemporal dementia, supranuclear gaze palsy, and chorea. Mov Disord. 2009;24: 1843–1847. doi:10.1002/mds.22701

109. Borroni B, Archetti S, Del Bo R, Papetti a, Buratti E, Bonvicini C, et al. TARDBP mutations in frontotemporal lobar degeneration: frequency, clinical features, and disease course. Rejuvenation Res. 2010;13: 509–517. doi:10.1089/rej.2010.1017

110. Chiò A, Calvo A, Moglia C, Restagno G, Ossola I, Brunetti M, et al. Amyotrophic lateral sclerosis-frontotemporal lobar dementia in 3 families with p.Ala382Thr TARDBP mutations. Arch Neurol. 2010;67: 1002–1009. doi:10.1001/archneurol.2010.173

111. Chiò A, Borghero G, Pugliatti M, Ticca A, Calvo A, Moglia C, et al. Large proportion of amyotrophic lateral sclerosis cases in Sardinia due to a single founder mutation of the TARDBP gene. Arch Neurol. 2011;68: 594–598. doi:10.1001/archneurol.2010.352

112. Kwiatkowski T, Bosco D, Leclerc A, Tamrazian E, Vanderburg C, Russ C, et al. Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. Science (80-). 2009;323: 1205–1208. doi:323/5918/1205 [pii]10.1126/science.1166066

113. Vance C, Rogelj B, Hortobagyi T, De Vos KJ, Nishimura AL, Sreedharan J, et al. Mutations in FUS, an RNA Processing Protein, Cause Familial Amyotrophic Lateral Sclerosis Type 6. Science (80-). 2009;323: 1208–1211. doi:10.1126/science.1165942

114. Van Langenhove T, Van Der Zee J, Sleegers K, Engelborghs S, Vandenberghe R, Gijselinck I, et al. Genetic contribution of FUS to frontotemporal lobar degeneration. Neurology. 2010;74: 366–371. doi:10.1212/WNL.0b013e3181ccc732

115. Huey ED, Ferrari R, Moreno JH, Jensen C, Morris CM, Potocnik F, et al. FUS and TDP43 genetic variability in FTD and CBS. Neurobiol Aging. 2012;33. doi:10.1016/j.neurobiolaging.2011.08.004

116. Hardy J, Rogaeva E. Motor neuron disease and frontotemporal dementia: Sometimes related, sometimes not. Experimental Neurology. 2014. pp. 75–83. doi:10.1016/j.expneurol.2013.11.006

117. Renton AE, Majounie E, Waite A, Simón-Sánchez J, Rollinson S, Gibbs JR, et al. A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. Neuron. Neuromuscular Diseases Research Unit, Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD 20892, USA.; 2011;72: 257–268. Available: http://www.sciencedirect.com/science/article/pii/S0896627311007975

118. Gijselinck I, Van Langenhove T, van der Zee J, Sleegers K, Philtjens S, Kleinberger G, et al. A C9orf72 promoter repeat expansion in a Flanders-Belgian cohort with disorders of the frontotemporal lobar degeneration-amyotrophic lateral sclerosis

spectrum: A gene identification study. Lancet Neurol. 2012;11: 54–65. doi:10.1016/S1474-4422(11)70261-7

119. Van Mossevelde S, van der Zee J, Cruts M, Van Broeckhoven C. Relationship between C9orf72 repeat size and clinical phenotype. Current Opinion in Genetics and Development. 2017. pp. 117–124. doi:10.1016/j.gde.2017.02.008

120. Ng ASL, Tan E-K. Intermediate C9orf72 alleles in neurological disorders: does size really matter? J Med Genet. 2017; Available: http://jmg.bmj.com/content/early/2017/07/07/jmedgenet-2017-104752.abstract

121. Nuytemans K, Bademci G, Kohli MM, Beecham GW, Wang L, Young JI, et al. C9orf72 intermediate repeat copies are a significant risk factor for parkinson disease. Ann Hum Genet. 2013;77: 351–363. doi:10.1111/ahg.12033

122. Majounie E, Renton AE, Mok K, Dopper EGP, Waite A, Rollinson S, et al. Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: A cross-sectional study. Lancet Neurol. 2012;11: 323–330. doi:10.1016/S1474-4422(12)70043-1

123. Yokoyama JS, Sirkis DW, Miller BL. C9ORF72 hexanucleotide repeats in behavioral and motor neuron disease: clinical heterogeneity and pathological diversity. Am J Neurodegener Dis. 2014;3: 1–18. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3986607&tool=pmcentrez&rendertype=abstract

124. Lesage S, Le Ber I, Condroyer C, Broussolle E, Gabelle A, Thobois S, et al. C9orf72 repeat expansions are a rare genetic cause of parkinsonism. Brain. 2013;136: 385–391. doi:10.1093/brain/aws357

125. Kohli MA, John-Williams K, Rajbhandary R, Naj A, Whitehead P, Hamilton K, et al. Repeat expansions in the C9ORF72 gene contribute to Alzheimer's disease in Caucasians. Neurobiol Aging. 2013;34. doi:10.1016/j.neurobiolaging.2012.10.003

126. Harms M, Benitez BA, Cairns N, Cooper B, Cooper P, Mayo K, et al. C9orf72 hexanucleotide repeat expansions in clinical Alzheimer disease. JAMA Neurol. 2013;70: 736–41. doi:10.1001/2013.jamaneurol.537

127. Cacace R, Van Cauwenberghe C, Bettens K, Gijselinck I, van der Zee J, Engelborghs S, et al. C9orf72 G4C2 repeat expansions in Alzheimer's disease and mild cognitive impairment. Neurobiol Aging. 2013;34. doi:10.1016/j.neurobiolaging.2012.12.019

128. Cooper-Knock J, Shaw PJ, Kirby J. The widening spectrum of C9ORF72-related disease; Genotype/phenotype correlations and potential modifiers of clinical phenotype. Acta Neuropathologica. 2014. pp. 333–345. doi:10.1007/s00401-014-1251-9

129. Ling SC, Polymenidou M, Cleveland DW. Converging mechanisms in als and FTD: Disrupted RNA and protein homeostasis. Neuron. 2013. pp. 416–438. doi:10.1016/j.neuron.2013.07.033

130. Yokoyama JS, Sirkis DW, Miller BL. C9ORF72 hexanucleotide repeats in behavioral and motor neuron disease: clinical heterogeneity and pathological diversity.

Am J Neurodegener Dis. 2014;3: 1–18. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3986607&tool=pmcentrez&rendertype=abstract

131. Jiang J, Zhu Q, Gendron TF, Saberi S, McAlonis-Downes M, Seelman A, et al. Gain of Toxicity from ALS/FTD-Linked Repeat Expansions in C9ORF72 Is Alleviated by Antisense Oligonucleotides Targeting GGGGCC-Containing RNAs. Neuron. 2016;90: 535–550. doi:10.1016/j.neuron.2016.04.006

132. Watts GDJ, Wymer J, Kovach MJ, Mehta SG, Mumm S, Darvish D, et al. Inclusion body myopathy associated with Paget disease of bone and frontotemporal dementia is caused by mutant valosin-containing protein. Nat Genet. 2004;36: 377–381. doi:10.1038/ng1332

133. Watts GDG, Thomasova D, Ramdeen SK, Fulchiero EC, Mehta SG, Drachman DA, et al. Novel VCP mutations in inclusion body myopathy associated with Paget disease of bone and frontotemporal dementia. Clin Genet. 2007;72: 420–426. doi:10.1111/j.1399-0004.2007.00887.x

134. Kimonis VE, Fulchiero E, Vesa J, Watts G. VCP disease associated with myopathy, Paget disease of bone and frontotemporal dementia: Review of a unique disorder. Biochimica et Biophysica Acta - Molecular Basis of Disease. 2008. pp. 744–748. doi:10.1016/j.bbadis.2008.09.003

135. Shaw CE. Capturing VCP: Another Molecular Piece in the ALS Jigsaw Puzzle. Neuron. 2010. pp. 812–814. doi:10.1016/j.neuron.2010.11.040

136. Johnson JO, Mandrioli J, Benatar M, Abramzon Y, Van Deerlin VM, Trojanowski JQ, et al. Exome Sequencing Reveals VCP Mutations as a Cause of Familial ALS. Neuron. 2010;68: 857–864. doi:10.1016/j.neuron.2010.11.036

137. DeJesus-Hernandez M, Desaro P, Johnston A, Ross OA, Wszolek ZK, Ertekin-Taner N, et al. Novel p.Ile151Val mutation in VCP in a patient of African American descent with sporadic ALS. Neurology. 2011;77: 1102–1103.

138. Koppers M, van Blitterswijk MM, Vlam L, Rowicka PA, van Vught PWJ, Groen EJN, et al. VCP mutations in familial and sporadic amyotrophic lateral sclerosis. Neurobiol Aging. 2012;33. doi:10.1016/j.neurobiolaging.2011.10.006

139. Abramzon Y, Johnson JO, Scholz SW, Taylor JP, Brunetti M, Calvo A, et al. Valosin-containing protein (VCP) mutations in sporadic amyotrophic lateral sclerosis. Neurobiol Aging. 2012;33. doi:10.1016/j.neurobiolaging.2012.04.005

140. Deng H-X, Chen W, Hong S-T, Boycott KM, Gorrie GH, Siddique N, et al. Mutations in UBQLN2 cause dominant X-linked juvenile and adult-onset ALS and ALS/dementia. Nature. 2011;477: 211–5. doi:10.1038/nature10353

141. Synofzik M, Maetzler W, Grehl T, Prudlo J, vom Hagen JM, Haack T, et al. Screening in ALS and FTD patients reveals 3 novel UBQLN2 mutations outside the PXX domain and a pure FTD phenotype. Neurobiol Aging. 2012;33. doi:10.1016/j.neurobiolaging.2012.07.002

142. Daoud H, Suhail H, Szuto A, Camu W, Salachas F, Meininger V, et al. UBQLN2 mutations are rare in French and French-Canadian amyotrophic lateral sclerosis. Neurobiol Aging. 2012;33. doi:10.1016/j.neurobiolaging.2012.03.015

143. van Rheenen W, van Blitterswijk M, Schellevis RD, Schelhaas HJ, de Visser M, van der Kooi AJ, et al. UBQLN2 in familial amyotrophic lateral sclerosis in The Netherlands. Neurobiol Aging. 2012;33: 2233.e7-2233.e8. Available: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=22676852

144. Kim HJ, Kwon MJ, Choi WJ, Oh KW, Oh S il, Ki CS, et al. Mutations in UBQLN2 and SIGMAR1 genes are rare in Korean patients with amyotrophic lateral sclerosis. Neurobiol Aging. 2014;35. doi:10.1016/j.neurobiolaging.2014.03.001

145. Fecto F, Yan J, Vemula SP, Liu E, Yang Y, Chen W, et al. SQSTM1 mutations in familial and sporadic amyotrophic lateral sclerosis. Arch Neurol. 2011;68: 1440–1446. doi:10.1001/archneurol.2011.250

146. Rubino E, Rainero I, Chio a., Rogaeva E, Galimberti D, Fenoglio P, et al. SQSTM1 mutations in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Neurology. 2012;79: 1556–1562. doi:10.1212/WNL.0b013e31826e25df

147. Shimizu H, Toyoshima Y, Shiga A, Yokoseki A, Arakawa K, Sekine Y, et al. Sporadic ALS with compound heterozygous mutations in the SQSTM1 gene. Acta Neuropathol. 2013;126: 453–459. doi:10.1007/s00401-013-1150-5

148. Hirano M, Nakamura Y, Saigoh K, Sakamoto H, Ueno S, Isono C, et al. Mutations in the gene encoding p62 in Japanese patients with amyotrophic lateral sclerosis. Neurology. 2013;80: 458–463. doi:10.1212/WNL.0b013e31827f0fe5

149. Teyssou E, Takeda T, Lebon V, Boillée S, Doukouré B, Bataillon G, et al. Mutations in SQSTM1 encoding p62 in amyotrophic lateral sclerosis: Genetics and neuropathology. Acta Neuropathol. 2013;125: 511–522. doi:10.1007/s00401-013-1090-0

150. Chen Y, Zheng Z-Z, Chen X, Huang R, Yang Y, Yuan L, et al. SQSTM1 mutations in Han Chinese populations with sporadic amyotrophic lateral sclerosis. Neurobiology of Aging. 2014. doi:10.1016/j.neurobiolaging.2013.09.008

151. Le Ber I, Camuzat A, Guerreiro R, Bouya-Ahmed K, Bras J, Nicolas G, et al. *SQSTM1* Mutations in French Patients With Frontotemporal Dementia or Frontotemporal Dementia With Amyotrophic Lateral Sclerosis. JAMA Neurol. 2013;70: 1582–1588. doi:10.1001/jamaneurol.2013.3849

152. Goode A, Layfield R. Recent advances in understanding the molecular basis of Paget disease of bone. J Clin Pathol. 2010;63: 199–203. doi:10.1136/jcp.2009.064428

153. Maruyama H, Morino H, Ito H, Izumi Y, Kato H, Watanabe Y, et al. Mutations of optineurin in amyotrophic lateral sclerosis. Nature. Department of Epidemiology, Research Institute for Radiation Biology and Medicine, Hiroshima University, Hiroshima 734-8553, Japan.; 2010;465: 223–226. Available: http://www.nature.com/nature/journal/v465/n7295/full/nature08971.html

154. Chio A, Calvo A, Mazzini L, Cantello R, Mora G, Moglia C, et al. Extensive genetics of ALS : A population-based study in Italy. Neurology. 2012;79: 1983–1989. doi:10.1212/WNL.0b013e3182735d36

155. Pottier C, Bieniek KF, Finch N, van de Vorst M, Baker M, Perkersen R, et al. Whole-genome sequencing reveals important role for TBK1 and OPTN mutations in frontotemporal lobar degeneration without motor neuron disease. Acta Neuropathol. 2015;130: 77–92. doi:10.1007/s00401-015-1436-x

156. Freischmidt A, Wieland T, Richter B, Ruf W, Schaeffer V, Müller K, et al. Haploinsufficiency of TBK1 causes familial ALS and fronto-temporal dementia. Nat Neurosci. 2015;18: 631–6. doi:10.1038/nn.4000

157. Pottier C, Bieniek K, Finch N, van de Vorst M, Baker M, Perkersen R, et al. Whole-genome sequencing reveals important role for TBK1 and OPTN mutations in frontotemporal lobar degeneration without motor neuron disease. Acta Neuropathol. 2015;130: 77–92.

158. Gijselinck I, Van Mossevelde S, Van Der Zee J, Sieben A, Philtjens S, Heeman B, et al. Loss of TBK1 is a frequent cause of frontotemporal dementia in a Belgian cohort. Neurology. 2015;85: 2116–2125. doi:10.1212/WNL.0000000000002220

159. Borghero G, Pugliatti M, Marrosu F, Marrosu MG, Murru MR, Floris G, et al. TBK1 is associated with ALS and ALS-FTD in Sardinian patients. Neurobiol Aging. 2016; 1–5. doi:10.1016/j.neurobiolaging.2016.03.028

160. Bannwarth S, Ait-El-Mkadem S, Chaussenot A, Genin EC, Lacas-Gervais S, Fragaki K, et al. A mitochondrial origin for frontotemporal dementia and amyotrophic lateral sclerosis through CHCHD10 involvement. Brain. 2014;137: 2329–2345. doi:10.1093/brain/awu138

161. Bannwarth S, Ait-El-Mkadem S, Chaussenot A, Genin EC, Lacas-Gervais S, Fragaki K, et al. Reply: Two novel mutations in conserved codons indicate that CHCHD10 is a gene associated with motor neuron disease. Brain. 2014. p. e310. doi:10.1093/brain/awu228

162. Chaussenot A, Le Ber I, Ait-El-Mkadem S, Camuzat A, de Septenville A, Bannwarth S, et al. Screening of CHCHD10 in a French cohort confirms the involvement of this gene in frontotemporal dementia with amyotrophic lateral sclerosis patients. Neurobiol Aging. 2014;35: 2884.e1-2884.e4. doi:10.1016/j.neurobiolaging.2014.07.022

163. Johnson JO, Glynn SM, Gibbs JR, Nalls MA, Sabatelli M, Restagno G, et al. Mutations in the CHCHD10 gene are a common cause of familial amyotrophic lateral sclerosis. Brain. 2014;137: e311. doi:10.1093/brain/awu265

164. Muller K, Andersen P, Hubers A, Marroquin N, Volk A, Danzer K, et al. Two novel mutations in conserved codons indicate that CHCHD10 is a gene associated with motor neuron disease. Brain. 2014;137: 1–3.

165. Skibinski G, Parkinson NJ, Brown JM, Chakrabarti L, Lloyd SL, Hummerich H, et al. Mutations in the endosomal ESCRTIII-complex subunit CHMP2B in frontotemporal dementia. Nat Genet. 2005;37: 806–808. doi:10.1038/ng1609

166. Cannon A, Baker M, Boeve B, Josephs K, Knopman D, Petersen R, et al. CHMP2B mutations are not a common cause of frontotemporal lobar degeneration. Neurosci Lett. 2006;398: 83–84. doi:10.1016/j.neulet.2005.12.056

167. Parkinson N, Ince PG, Smith MO, Highley R, Skibinski G, Andersen PM, et al. ALS phenotypes with mutations in CHMP2B (charged multivesicular body protein 2B). Neurology. 2006;67: 1074–1077. doi:10.1212/01.wnl.0000231510.89311.8b

168. van der Zee J, Urwin H, Engelborghs S, Bruyland M, Vandenberghe R, Dermaut B, et al. CHMP2B C-truncating mutations in frontotemporal lobar degeneration are associated with an aberrant endosomal phenotype in vitro. Hum Mol Genet. 2008;17: 313–322. doi:10.1093/hmg/ddm309

169. Cox LE, Ferraiuolo L, Goodall EF, Heath PR, Higginbottom A, Mortiboys H, et al. Mutations in CHMP2B in lower motor neuron predominant amyotrophic lateral sclerosis (ALS). PLoS One. 2010;5. doi:10.1371/journal.pone.0009872

170. Ji A, Zhang X, Chen W, Huang W. Genetics insight into the amyotrophic lateral sclerosis/frontotemporal dementia spectrum. J Med Genet. 2017;0: 1–10.

171. Clark LN, Poorkaj P, Wszolek Z, Geschwind DH, Nasreddine ZS, Miller B, et al. Pathogenic implications of mutations in the tau gene in pallido-ponto-nigral degeneration and related neurodegenerative disorders linked to chromosome 17. Proc Natl Acad Sci U S A. 1998;95: 13103–7. doi:10.1073/pnas.95.22.13103

172. Hutton M, Lendon CL, Rizzu P, Baker M, Froelich S, Houlden H, et al. Association of missense and 5'-splice-site mutations in tau with the inherited dementia FTDP-17. Nature. 1998;393: 702–5. doi:10.1038/31508

173. Poorkaj P, Bird TD, Wijsman E, Nemens E, Garruto RM, Anderson L, et al. Tau is a candidate gene for chromosome 17 frontotemporal dementia. Ann Neurol. 1998;43: 815–25. doi:10.1002/ana.410430617

174. Spillantini MG, Murrell JR, Goedert M, Farlow MR, Klug A, Ghetti B. Mutation in the tau gene in familial multiple system tauopathy with presenile dementia. Proc Natl Acad Sci U S A. 1998;95: 7737–41. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=22742&tool=pmcentrez&rendertype=abstract

175. Cruts M, Gijselinck I, van der Zee J, Engelborghs S, Wils H, Pirici D, et al. Null mutations in progranulin cause ubiquitin-positive frontotemporal dementia linked to chromosome 17q21. Nature. 2006;442: 920–4. doi:10.1038/nature05017

176. Baker M, Mackenzie IR, Pickering-Brown SM, Gass J, Rademakers R, Lindholm C, et al. Mutations in progranulin cause tau-negative frontotemporal dementia linked to chromosome 17. Nature. 2006;442: 916–9. doi:10.1038/nature05016

177. van Swieten JC, Heutink P. Mutations in progranulin (GRN) within the spectrum of clinical and pathological phenotypes of frontotemporal dementia. The Lancet Neurology. 2008. pp. 965–974. doi:10.1016/S1474-4422(08)70194-7

178. Gass J, Prudencio M, Stetler C, Petrucelli L. Progranulin: An emerging target for FTLD therapies. Brain Research. 2012. pp. 118–128. doi:10.1016/j.brainres.2012.01.047

179. Perry DC, Lehmann M, Yokoyama JS, Karydas A, Lee JJ, Coppola G, et al. Progranulin Mutations as Risk Factors for Alzheimer Disease. JAMA Neurol. 2013;70: 774. doi:10.1001/2013.jamaneurol.393

180. Naasan G, Rabinovici GD, Ghosh P, Elofson JD, Miller BL, Coppola G, et al. Amyloid in dementia associated with familial FTLD: not an innocent bystander. Neurocase. 2015; 1–8. doi:10.1080/13554794.2015.1046458

181. Van Deerlin VM, Sleiman PM a, Martinez-Lage M, Chen-Plotkin A, Wang L-S, Graff-Radford NR, et al. Common variants at 7p21 are associated with frontotemporal lobar degeneration with TDP-43 inclusions. Nat Genet. Nature Publishing Group; 2010;42: 234–9. doi:10.1038/ng.536

182. Finch N, Carrasquillo MM, Baker M, Rutherford NJ, Coppola G, Dejesus-Hernandez M, et al. TMEM106B regulates progranulin levels and the penetrance of FTLD in GRN mutation carriers. Neurology. From the Departments of Neuroscience (N.F., M.M.C., M.B., N.J.R., M.D.-H., R.C., T.H., J.C., J.G., S.G.Y., D.W.D., R.R.) and Neurology (R.U., Z.K.W., N.R.G.-R.), Mayo Clinic, Jacksonville, FL; Department of Neurology (G.C., D.H.G.), The David Geffen Schoo; 2010;76: 467–474. Available: http://www.neurology.org/content/early/2010/12/22/WNL.0b013e31820a0e3b.1.long

183. Cruchaga C, Graff C, Chiang H-H, Wang J, Hinrichs AL, Spiegel N, et al. Association of TMEM106B gene polymorphism with age at onset in granulin mutation carriers and plasma granulin protein levels. Arch Neurol. 2011;68: 581–6. doi:10.1001/archneurol.2010.350

184. van Blitterswijk M, Mullen B, Nicholson AM, Bieniek KF, Heckman MG, Baker MC, et al. TMEM106B protects C9ORF72 expansion carriers against frontotemporal dementia. Acta Neuropathol. 2014;127: 397–406. doi:10.1007/s00401-013-1240-4

185. Gallagher MD, Suh E, Grossman M, Elman L, McCluskey L, Van Swieten JC, et al. TMEM106B is a genetic modifier of frontotemporal lobar degeneration with C9orf72 hexanucleotide repeat expansions. Acta Neuropathol. 2014;127: 407–18. doi:10.1007/s00401-013-1239-x

186. Vass R, Ashbridge E, Geser F, Hu WT, Grossman M, Clay-Falcone D, et al. Risk genotypes at TMEM106B are associated with cognitive impairment in amyotrophic lateral sclerosis. Acta Neuropathol. Department of Neurology, University of Pennsylvania School of Medicine, 3 W Gates, 3400 Spruce Street, Philadelphia, PA, 19104, USA.; 2010; Available: http://www.springerlink.com/content/u8m7j3511141p451/

187. Ferrari R, Hernandez DG, Nalls MA, Rohrer JD, Ramasamy A, Kwok JBJ, et al. Frontotemporal dementia and its subtypes: A genome-wide association study. Lancet Neurol. 2014;13: 686–699. doi:10.1016/S1474-4422(14)70065-1

188. Alzheimer's Association. 2016 Alzheimer's Disease Facts and Figures. In: Alzheimer's and Dementia [Internet]. 2016 [cited 7 Feb 2017] pp. 1–84. Available: http://www.alz.org/documents_custom/2016-facts-and-figures.pdf

189. Naj AC, Schellenberg GD, for the Alzheimer's Disease Genetics Consortium (ADGC). Genomic variants, genes, and pathways of Alzheimer's disease: An overview. Am J Med Genet Part B Neuropsychiatr Genet. 2017;174: 5–26. doi:10.1002/ajmg.b.32499

190. Guerreiro RJ, Gustafson DR, Hardy J. The genetic architecture of Alzheimer's disease: Beyond APP, PSENS and APOE. Neurobiology of Aging. 2012. pp. 437–456. doi:10.1016/j.neurobiolaging.2010.03.025

191. Clayton D, Chandra V, Fratiglioni L, Graves AB, Heyman A, Jorm AF, et al. Familial aggregation of Alzheimer's disease and related disorders: a collaborative re-analysis of case-control studies. Int J Epidemiol. 1991;20 Suppl 2: S13--20. Available: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=med3&NEWS=N&AN=1833349

192. Bird TD. Early-Onset Familial Alzheimer Disease. Gene Reviews. 2012.

193. Hardy JA, Higgins GA. Alzheimer's disease: the amyloid cascade hypothesis. Science (80-). 1992;256: 184–185. doi:10.1126/science.1566067

194. Goate A, Chartier-Harlin MC, Mullan M, Brown J, Crawford F, Fidani L, et al. Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. Nature. 1991;349: 704–6. doi:10.1038/349704a0

195. Hardy J. Molecular classifications of Alzheimer's disease. Lancet. 1991;1: 1342–1343.

196. Chartier-Harlin M, Crawford F, H H, A W, D H, L F, et al. Early-onset Alzheimer's disease caused by mutations at codon 717 of the beta-amyloid precursor protein gene. Nature. 1991;353: 844–846. Available: papers2://publication/uuid/35530EF8-1C8F-48CF-B133-065FC906C22C

197. Mullan M, Crawford F, Axelman K, Houlden H, Lilius L, Winblad B, et al. A pathogenic mutation for probable Alzheimer's disease in the APP gene at the N-terminus of beta-amyloid. Nat Genet. 1992;1: 345–347. doi:10.1038/ng0892-345

198. Fidani L, Rooke K, Chartier-harlin M-C, Hughes D, Tanzi R, Mullan M, et al. Screening for mutations in the open reading frame and promoter of the β-amyloid precursor protein gene in familial alzheimer's disease: Identification of a further family with APP717 val→Ile. Hum Mol Genet. 1992;1. doi:10.1093/hmg/1.3.165

199. Hardy J. Amyloid, the presenilins and Alzheimer's disease. Trends in Neurosciences. 1997. pp. 154–159. doi:10.1016/S0166-2236(96)01030-2

200. Cruchaga C, Chakraverty S, Mayo K, Vallania FLM, Mitra RD, Faber K, et al. Rare variants in APP, PSEN1 and PSEN2 increase risk for AD in late-onset Alzheimer's disease families. PLoS One. 2012;7. doi:10.1371/journal.pone.0031039

201. Wisniewski KE, Wisniewski HM, Wen GY. Occurrence of neuropathological changes and dementia of Alzheimer's disease in Down's syndrome. Ann Neurol. 1985;17: 278–282. doi:10.1002/ana.410170310

202. Millan Sanchez M, Heyn SN, Das D, Moghadam S, Martin KJ, Salehi A. Neurobiological elements of cognitive dysfunction in down syndrome: Exploring the role of APP. Biological Psychiatry. 2012. pp. 403–409. doi:10.1016/j.biopsych.2011.08.016

203. Jonsson T, Atwal JK, Steinberg S, Snaedal J, Jonsson P V, Bjornsson S, et al. A mutation in APP protects against Alzheimer's disease and age-related cognitive decline. Nature. 2012;488: 1–4.

204. Bamne MN, Demirci FY, Berman S, Snitz BE, Rosenthal SL, Wang X, et al. Investigation of an amyloid precursor protein protective mutation (A673T) in a North American case-control sample of late-onset Alzheimer's disease. Neurobiol Aging. 2014;35. doi:10.1016/j.neurobiolaging.2014.01.020

205. Ting SKS, Chong M-S, Kandiah N, Hameed S, Tan L, Au W-L, et al. Absence of A673T amyloid-beta precursor protein variant in Alzheimer's disease and other neurological diseases. Neurobiol Aging. 2013;34: 2441.e7-8. doi:10.1016/j.neurobiolaging.2013.04.012

206. O'Brien RJ, Wong PC. Amyloid Precursor Protein Processing and Alzheimer's Disease. Annu Rev Neurosci. 2011; 183–202. doi:10.1146/annurev-neuro-061010-113613

207. Rogaev EI, Sherrington R, Rogaeva E a, Levesque G, Ikeda M, Liang Y, et al. Familial Alzheimer's disease in kindreds with missense mutations in a gene on chromosome 1 related to the Alzheimer's disease type 3 gene. Nature. 1995. pp. 775–778. doi:10.1038/376775a0

208. Sherrington R, Rogaev EI, Liang Y, Rogaeva EA, Levesque G, Ikeda M, et al. Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. Nature. 1995;375: 754–60. doi:10.1038/375754a0

209. Cruts M, Theuns J, Van Broeckhoven C. Locus-specific mutation databases for neurodegenerative brain diseases. Hum Mutat. 2012;33: 1340–1344. doi:10.1002/humu.22117

210. M C, J T, C VB. Alzheimer Disease & Frontotemporal Dementia Mutation Database [Internet]. [cited 1 Feb 2017]. Available: http://www.molgen.vib-ua.be/ADMutations

211. Van Cauwenberghe C, Van Broeckhoven C, Sleegers K. The genetic landscape of Alzheimer disease: clinical implications and perspectives. Genet Med. 2016;18: 421–430. doi:10.1038/gim.2015.117

212. Bateman RJ, Xiong C, Benzinger TLS, Fagan AM, Goate A, Fox NC, et al. Clinical and biomarker changes in dominantly inherited Alzheimer's disease. N Engl J Med. 2012;367: 795–804. doi:10.1056/NEJMoa1202753

213. Pericak-Vance MA, Yamaoka LH, Haynes CS, Speer MC, Haines JL, Gaskell PC, et al. Genetic linkage studies in Alzheimer's disease families. Exp Neurol. 1988;102: 271–279. Available: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=3197787

214. Strittmatter WJ, Saunders AM, Schmechel D, Pericak-Vance M, Enghild J, Salvesen GS, et al. Apolipoprotein E: high-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease. Proc Natl Acad Sci U S A. 1993;90: 1977–81. doi:10.1073/pnas.90.5.1977

215. Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GW, et al. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. Sci (New York, NY). Department of Medicine, Joseph and Kathleen Bryan Alzheimer's Disease Research Center, Duke University Medical Center, Durham, NC 27710.; 1993;261: 921–923. Available: papers://f5e128fc-29df-4fa6-b2d2-ee5f59b2d8d6/Paper/p655

216. Saunders AM, Strittmatter WJ, Schmechel D, George-Hyslop PH, Pericak-Vance MA, Joo SH, et al. Association of apolipoprotein E allele epsilon 4 with late-onset familial and sporadic Alzheimer's disease. Neurology. Department of Medicine [Neurology], Joseph and Kathleen Bryan Alzheimer's Disease Research Center, Duke University Medical Center, Durham, NC.; 1993;43: 1467–1472. Available: http://www.ncbi.nlm.nih.gov/sites/entrez?Db=pubmed&Cmd=Retrieve&list_uids=8350998&dopt=abstractplus

217. Farrer LA, Cupples LA, Haines JL, Hyman B, Kukull WA, Mayeux R, et al. Effects of Age, Sex, and Ethnicity on the Association Between Apolipoprotein E Genotype and Alzheimer Disease: A Meta-analysis. JAMA: The Journal of the American Medical Association. 1997. pp. 1349–1356. doi:10.1001/jama.1997.03550160069041

218. Eisenberg DTA, Kuzawa CW, Hayes MG. Worldwide allele frequencies of the human apolipoprotein E gene: Climate, local adaptations, and evolutionary history. Am J Phys Anthropol. 2010;143: 100–111. doi:10.1002/ajpa.21298

219. Raber J, Huang Y, Ashford JW. ApoE genotype accounts for the vast majority of AD risk and AD pathology. Neurobiol Aging. Departments of Behavioral Neuroscience and Neurology, L470, ONPRC, Oregon Health and Science University, 3181 S.W. Sam Jackson Park Road, Portland, OR 97201, USA. raberj@ohsu.edu; 2004;25: 641–650. Available: http://www.sciencedirect.com/science?_ob=ArticleURL&_udi=B6T09-4CBW8D5-1&_user=4430&_coverDate=06/30/2004&_rdoc=1&_fmt=high&_orig=search&_origin=search&_sort=d&_docanchor=&view=c&_acct=C000059594&_version=1&_urlVersion=0&_userid=4430&md5=26d422ca29e361c3c512ad44f04dc168&searchtype=a

220. Crean S, Ward A, Mercaldi CJ, Collins JM, Cook MN, Baker NL, et al. Apolipoprotein E ε4 prevalence in Alzheimer's disease patients varies across global populations: a systematic literature review and meta-analysis. Dement Geriatr Cogn Disord. United BioSource Corporation, Lexington, MA 02420, USA. sheila.crean@unitedbiosource.com; 2011;31: 20–30. Available: http://content.karger.com/produktedb/produkte.asp?typ=fulltext&file=000321984

221. Huang Y, Mucke L. Alzheimer mechanisms and therapeutic strategies. Cell. 2012. pp. 1204–1222. doi:10.1016/j.cell.2012.02.040

222. Huang Y. Abeta-independent roles of apolipoprotein E4 in the pathogenesis of Alzheimer's disease. Trends Mol Med. The Gladstone Institute of Neurological Disease,

University of California, San Francisco, CA 94158, USA. yhuang@gladstone.ucsf.edu; 2010;16: 287–294. Available: http://www.sciencedirect.com/science?_ob=ArticleURL&_udi=B6W7J-505NCV7-1&_user=4430&_coverDate=06/30/2010&_rdoc=1&_fmt=high&_orig=search&_origin=search&_sort=d&_docanchor=&view=c&_acct=C000059594&_version=1&_urlVersion=0&_userid=4430&md5=b8eb892971381b6928f6821b6221c7b9&searchtype=a

223. Mahley RW, Huang Y. Apolipoprotein E Sets the Stage: Response to Injury Triggers Neuropathology. Neuron. 2012. pp. 871–885. doi:10.1016/j.neuron.2012.11.020

224. Pastor P, Roe CM, Villegas A, Bedoya G, Chakraverty S, García G, et al. Apolipoprotein Eε4 modifies Alzheimer's disease onset in an E280A PS1 kindred. Ann Neurol. 2003;54: 163–169. doi:10.1002/ana.10636

225. Wijsman EM, Daw EW, Yu X, Steinbart EJ, Nochlin D, Bird TB, et al. APOE and other loci affect age-at-onset in Alzheimer's disease families with PS2 mutation. Am J Med Genet - Neuropsychiatr Genet. 2005;132 B: 14–20. doi:10.1002/ajmg.b.30087

226. Guerreiro R, Wojtas A, Bras J, Carrasquillo M, Rogaeva E, Majounie E, et al. *TREM2* Variants in Alzheimer's Disease. N Engl J Med. 2013;368: 117–127. doi:10.1056/NEJMoa1211851

227. Jonsson T, Stefansson H, Steinberg S, Jonsdottir I, Jonsson P V, Snaedal J, et al. Variant of TREM2 associated with the risk of Alzheimer's disease. N Engl J Med. 2013;368: 107–116. doi:10.1056/NEJMoa1211103

228. Korvatska O, Leverenz JB, Jayadev S, McMillan P, Kurtz I, Guo X, et al. R47H Variant of TREM2 Associated With Alzheimer Disease in a Large Late-Onset Family: Clinical, Genetic, and Neuropathological Study. JAMA Neurol. 2015;72: 920–7. doi:10.1001/jamaneurol.2015.0979

229. Nicolas G, Charbonnier C, Campion D. From common to rare variants: The genetic component of Alzheimer disease. Human Heredity. 2017. pp. 129–141. doi:10.1159/000452256

230. Guerreiro RJ, Lohmann E, Brás JM, Gibbs JR, Rohrer JD, Gurunlian N, et al. Using exome sequencing to reveal mutations in TREM2 presenting as a frontotemporal dementia-like syndrome without bone involvement. JAMA Neurol. 2013;70: 78–84. doi:10.1001/jamaneurol.2013.579

231. Bianchin MM, Capella HM, Chaves DL, Steindel M, Grisard EC, Ganev GG, et al. Nasu-Hakola Disease (Polycystic Lipomembranous Osteodysplasia with Sclerosing Leukoencephalopathy-PLOSL): A Dementia Associated with Bone Cystic Lesions. From Clinical to Genetic and Molecular Aspects. Cellular and Molecular Neurobiology. 2004. pp. 1–24. doi:10.1023/B:CEMN.0000012721.08168.ee

232. Guerreiro RJ, Lohmann E, Brás JM, Gibbs JR, Rohrer JD, Gurunlian N, et al. Using Exome Sequencing to Reveal Mutations in TREM2 Presenting as a Frontotemporal Dementia-like Syndrome Without Bone Involvement. Arch Neurol. 2012; 1–7.

233. Grupe A, Abraham R, Li Y, Rowland C, Hollingworth P, Morgan A, et al. Evidence for novel susceptibility genes for late-onset Alzheimer's disease from a genome-wide association study of putative functional variants. Hum Mol Genet. 2007;16: 865–873. doi:10.1093/hmg/ddm031

234. Shen L, Jia J. An Overview of Genome-Wide Association Studies in Alzheimer???s Disease. Neuroscience Bulletin. 2016. pp. 183–190. doi:10.1007/s12264-016-0011-3

235. Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet. 2013;45: 1452–8. doi:10.1038/ng.2802

236. Jones L, Holmans PA, Hamshere ML, Harold D, Moskvina V, Ivanov D, et al. Genetic evidence implicates the immune system and cholesterol metabolism in the aetiology of Alzheimer's disease. PLoS One. 2010;5: e13950. doi:10.1371/journal.pone.0013950

237. Steele NZ, Carr JS, Bonham LW, Geier EG, Damotte V, Miller ZA, et al. Fine-mapping of the human leukocyte antigen (HLA) locus as a risk factor for Alzheimer's disease: A Case-Control Study. Plos Medicine. 2017: Under revisions.

238. Heneka MT, Golenbock DT, Latz E. Innate immunity in Alzheimer's disease. Nat Immunol. 2015;16: 229–236. doi:10.1038/ni.3102

239. Briones N, Dinu V. Data mining of high density genomic variant data for prediction of Alzheimer's disease risk. BMC Med Genet. 2012;13: 7. doi:10.1186/1471-2350-13-7

240. Lacour A, Espinosa A, Louwersheimer E, Heilmann S, Hernández I, Wolfsgruber S, et al. Genome-wide significant risk factors for Alzheimer's disease: role in progression to dementia due to Alzheimer's disease among subjects with mild cognitive impairment. Mol Psychiatry. 2016; 1–8. doi:10.1038/mp.2016.18

241. Chouraki V, Reitz C, Maury F, Bis JC, Bellenguez C, Yu L, et al. Evaluation of a Genetic Risk Score to Improve Risk Prediction for Alzheimer's Disease. J Alzheimers Dis. 2016;53: 921–32. doi:10.3233/JAD-150749

242. Escott-Price V, Shoai M, Pither R, Williams J, Hardy J. Polygenic score prediction captures nearly all common genetic risk for Alzheimer's disease. Neurobiol Aging. 2016;49: 214.e7-214.e11. doi:10.1016/j.neurobiolaging.2016.07.018

243. Escott-Price V, Myers AJ, Huentelman M, Hardy J. Polygenic Risk Score Analysis of Pathologically Confirmed Alzheimer's Disease. Ann Neurol. 2017; n/a--n/a. doi:10.1002/ana.24999

244. Tosto G, Bird TD, Tsuang D, Bennett DA, Boeve BF, Cruchaga C, et al. Polygenic risk scores in familial Alzheimer disease. Neurology. 2017;88: 1180–1186. doi:10.1212/WNL.0000000000003734

245. Desikan RS, Fan CC, Wang Y, Schork AJ, Cabral HJ, Cupples LA, et al. Genetic assessment of age-associated Alzheimer disease risk: Development and validation of a polygenic hazard score. PLOS Med. Public Library of Science; 2017;14: 1–17. doi:10.1371/journal.pmed.1002258

246. Le Couteur DG, Hunter S, Brayne C. Solanezumab and the amyloid hypothesis for Alzheimer's disease. BMJ. 2016;355.

247. Hardy J. The amyloid hypothesis for Alzheimer's disease: A critical reappraisal. Journal of Neurochemistry. 2009. pp. 1129–1134. doi:10.1111/j.1471-4159.2009.06181.x

248. Jack CR, Knopman DS, Jagust WJ, Petersen RC, Weiner MW, Aisen PS, et al. Tracking pathophysiological processes in Alzheimer's disease: An updated hypothetical model of dynamic biomarkers. The Lancet Neurology. 2013. pp. 207–216. doi:10.1016/S1474-4422(12)70291-0

249. Palop JJ, Mucke L. Network abnormalities and interneuron dysfunction in Alzheimer disease. Nat Rev Neurosci. 2016;17: 777–792. doi:10.1038/nrn.2016.141

250. Vossel KA, Beagle AJ, Rabinovici GD, Shu H, Lee SE, Naasan G, et al. Seizures and epileptiform activity in the early stages of Alzheimer disease. JAMA Neurol. 2013;70: 1158–66. doi:10.1001/jamaneurol.2013.136

251. Chung W-S, Welsh CA, Barres BA, Stevens B. Do glia drive synaptic and cognitive impairment in disease? Nat Neurosci. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.; 2015;18: 1539–1545. Available: http://dx.doi.org/10.1038/nn.4142

252. Dubal DB, Zhu L, Sanchez PE, Worden K, Broestl L, Johnson E, et al. Life Extension Factor Klotho Prevents Mortality and Enhances Cognition in hAPP Transgenic Mice. J Neurosci. 2015;35: 2358–2371. doi:10.1523/JNEUROSCI.5791-12.2015

253. Malinow R. New developments on the role of NMDA receptors in Alzheimer's disease. Current Opinion in Neurobiology. 2012. pp. 559–563. doi:10.1016/j.conb.2011.09.001

254. Bliss TVP, Collingridge GL, Morris RGM. Synaptic plasticity in health and disease: introduction and overview. Philos Trans R Soc Lond B Biol Sci. 2014;369: 20130129. doi:10.1098/rstb.2013.0129

255. Yokoyama JS, Wang Y, Schork AJ, Thompson WK, Karch CM, Cruchaga C, et al. Association between genetic traits for immune-mediated diseases and Alzheimer disease. JAMA Neurol. 2016;73. doi:10.1001/jamaneurol.2016.0150

256. Witoelar A, Jansen IE, Wang Y, Desikan RS, Raphael Gibbs J, Blauwendraat C, et al. Genome-wide Pleiotropy Between Parkinson Disease and Autoimmune Diseases. JAMA Neurol. 2017; doi:10.1001/jamaneurol.2017.0469

257. Atkin G, Paulson H. Ubiquitin pathways in neurodegenerative disease. Front Mol Neurosci. 2014;7: 63. doi:10.3389/fnmol.2014.00063

258. Barnes DE, Yaffe K. The projected effect of risk factor reduction on Alzheimer's disease prevalence. Lancet Neurol. Department of Psychiatry, University of California, San Francisco, San Francisco, CA 94121, USA. deborah.barnes@ucsf.edu; 2011;10: 819–828. Available: papers://f5e128fc-29df-4fa6-b2d2-ee5f59b2d8d6/Paper/p2178

259. Norton S, Matthews FE, Barnes DE, Yaffe K, Brayne C. Potential for primary prevention of Alzheimer's disease: An analysis of population-based data. Lancet Neurol. 2014;13: 788–794. doi:10.1016/S1474-4422(14)70136-X

260. Zeldich E, Chen C Di, Colvin TA, Bove-Fenderson EA, Liang J, Tucker Zhou TB, et al. The neuroprotective effect of Klotho is mediated via regulation of members of the redox system. J Biol Chem. 2014;289: 24700–24715. doi:10.1074/jbc.M114.567321

261. Yokoyama JS, Evans DS, Coppola G, Kramer JH, Tranah GJ, Yaffe K. Genetic modifiers of cognitive maintenance among older adults. Hum Brain Mapp. 2014;35: 4556–4565. doi:10.1002/hbm.22494

262. Dubal DB, Yokoyama JS, Zhu L, Broestl L, Worden K, Wang D, et al. Life extension factor klotho enhances cognition. Cell Rep. The Authors; 2014;7: 1065–1076. doi:10.1016/j.celrep.2014.03.076

263. Yokoyama JS, Sturm VE, Bonham LW, Klein E, Arfanakis K, Yu L, et al. Variation in longevity gene KLOTHO is associated with greater cortical volumes. Ann Clin Transl Neurol. 2015;2: 215–30. doi:10.1002/acn3.161

264. Yokoyama JS, Marx G, Brown J, Bonham LW, Wang D, Coppola G, et al. Systemic klotho is associated with KLOTHO variation and predicts intrinsic cortical connectivity in healthy human aging. Brain Imaging Behav. 2016; In Press.

265. Abraham CR, Mullen PC, Tucker-Zhou T, Chen CD, Zeldich E. Klotho Is a Neuroprotective and Cognition-Enhancing Protein. Vitamins and Hormones. 2016. pp. 215–238. doi:10.1016/bs.vh.2016.02.004

266. Harris SE, Deary IJ. The genetics of cognitive ability and cognitive ageing in healthy older people. Trends in Cognitive Sciences. 2011. pp. 388–394. doi:10.1016/j.tics.2011.07.004

267. Brooks-Wilson AR. Genetics of healthy aging and longevity. Human Genetics. 2013. pp. 1323–1338. doi:10.1007/s00439-013-1342-z

268. Vijg J, Suh Y. Genetics of longevity and aging. Annu Rev Med. 2005;56: 193–212. doi:10.1146/annurev.med.56.082103.104617

269. Nelson M, Tipney H, Painter J, Shen J, Nicoletti P, Shen Y, et al. The support of human genetic evidence for approved drug indications. Nat Genet. 2015;47: 856–862. doi:10.1038/ng.3314

Chapter II:

Fine-mapping of the human leukocyte antigen locus as a risk factor for Alzheimer's disease: A case–control study

Introduction

Alzheimer's disease (AD) is a progressive neurodegenerative disorder and has a global burden of approximately 46 million people worldwide, with prevalence projected to double over the next 20 years [1]. The hallmark features of the disease include the accumulation of amyloid plaques, tau neurofibrillary tangles, and neuronal destruction, leading to brain atrophy and loss of cognitive function. The etiology of these processes stems from synergistic interactions of environmental and genetic factors, many of which remain obscure and therefore complicate research efforts aimed at identifying efficacious therapies.

The three largest genetic contributors identified thus far are rare variants in amyloid precursor protein (*APP*) and presenilin 1 and 2 (*PSEN1, PSEN2*) [2]. These variants are uncommon, cause an early onset form of the disease, and typically segregate in an autosomal dominant fashion. Studies of late onset AD (typically defined as onset age >65 years) have demonstrated the risk of sequence variants such as the common ε4 allele of apolipoprotein E (*APOE*), rare variation in *TREM2* [3–7], and *MAPT* [8], as well as numerous common variants contributing modest AD risk [9], including single nucleotide polymorphisms (SNPs) in the following loci: *CR1, BIN1, INPP5D, MEF2C, CD2AP, ZCWPW1, NME8, EPHA1, CLU, PICALM, MS4A4, CELF1, FERMT2, ABCA7, CD33, CASS4, PTK2B, SORL1, SLC24A4-RIN3, DSG2*, and *HLA-DRB5/HLA-DRB1* [9,10]. However, there remain additional unexplained genetic contributions to non-familial forms of AD, suggesting polygenic contributors as well as the potential for epistatic and epigenetic interactions [11].

There is increasing support for the role of neuroinflammation in the etiology of AD as well as evidence that inflammatory processes are an early event in the brains of patients with AD [12]. Several studies have provided biochemical and histological evidence of classic immune components, including active microglia [13–15], complement factors [16,17], inflammatory cytokines [18], and C-reactive protein [19] within the parenchyma of AD brains. This is further supported by work in mouse models providing strong evidence for the role of complement-dependent destruction of synapses by phagocytic microglia prior to plaque deposition; similar mechanisms may even contribute to age-related cognitive decline [20,21]. Given these findings, there is a great deal of interest in identifying genetic determinants of inflammation related to AD susceptibility.

Located on chromosome 6p21, the major histocompatibility complex (MHC) is a dense region of approximately 150 genes that encode the human leukocyte antigen (HLA) immunoregulatory proteins [22]. Because of their proximity to each other, many of the MHC genes exist in linkage disequilibrium (LD) and are inherited as haplotypes with varying frequencies in global populations. MHC genes encode cell surface receptors and are classified based on their ability to present endogenous or exogenous antigens to T cells. MHC class I proteins exist on the surface of all nucleated cells and present fragments of antigens generated intracellularly to CD8+ T cells to induce a cytokine-mediated immune response. MHC class II molecules are only expressed by professional antigen-presenting cells, including B cells, macrophages, and microglia, and present exogenous material taken into the cell via endocytic vesicles to CD4+ T

cells. Together, the diverse repertoire of the human immune system partly stems from the extremely polymorphic nature of the MHC class I and II regions.

Many associations are established between neurodegenerative and autoimmune diseases, specific class I and II alleles, and combinations of alleles (haplotypes) in the HLA region. Previous genome-wide association studies (GWASs), pleiotropic analyses, and meta-analyses by our group and others have investigated MHC susceptibility loci in a wide range of diseases, including AD [9,23,24]. However, because of the complex genetic organization of the HLA region and differences in the haplotype substructure of different ethnic populations, as well as differences in sequencing and allelic imputation methods, studies have yet to definitively elucidate which genes and specific alleles contribute to the observed association signals.

As mentioned, *HLA-DRB5/HLA-DRB1* has been implicated in numerous GWASs as a significant contributor to AD risk [9]. This prior work has established a significant association of the HLA locus to AD risk in over 75,000 individuals, yet the specific allele or alleles contributing to this association remain elusive. We thus used a robust HLA imputation method and case–control approach to fine-map the contributions of HLA polymorphisms and haplotypes to AD in over 11,500 patients and controls from independent cohorts from the University of California, San Francisco (UCSF) Memory and Aging Center (MAC) and the Alzheimer's Disease Genetics Consortium (ADGC). We also examined longitudinal neuropsychological measures of cognitive function and cross-sectional biomarker data from cerebrospinal fluid (CSF) from the Alzheimer's Disease Neuroimaging Initiative (ADNI) to assess the clinical relevance of identified risk haplotypes.

Materials and methods

Participants were consented (as described below) for research in accordance with the Institutional Review Board at the University of California, San Francisco, and Institutional Review Boards at each site for multicenter study data approved all aspects of this study as they fall under the purview of the respective research groups (ADNI and ADGC).

Participants

UCSF MAC Cohort. The participants included in this study were 309 white individuals over the age of 50 years, including 191 controls and 118 individuals with AD seen at the UCSF MAC between 1999–2012 who were genotyped as part of their participation in longitudinal research on neurodegenerative disease and healthy cognitive aging. DNA from the UCSF MAC cohort was collected from 2000–2012, and genotyping was performed in 2012. Because individuals are followed up longitudinally, we verified clinical diagnosis at the beginning of this study (May 2015). A multidisciplinary team of neurologists, neuropsychologists, and nurses performed a detailed evaluation on individuals with AD and established a diagnosis according to consensus criteria for AD [25]. Individuals included as controls underwent a similar assessment and were diagnosed as having normal cognition for their age. Participants who carried a known genetic risk variant in *APP, PSEN1*, or *PSEN2* were excluded from this study. Participants or surrogates completed written informed consent for all genetic research related to neurodegenerative disease and healthy cognitive aging during their initial visit in accordance with the Institutional Review Board at the University of California, San Francisco.

ADGC. The ADGC is an NIH-funded collection of GWAS data created for the goal of identifying genetic contributions to late-onset AD. Participants included in this study were from 30 merged datasets combined by Boehme, Mukherjee, Crane, and Kauwe and included 28,730 individuals carrying either an AD or cognitively normal control clinical diagnosis [26] A list of the datasets and basic information is included in S1 Table; full details on the datasets and the merging process are available at http://kauwelab.byu.edu/Portals/22/adgc_combined_1000G_12032014.pdf [26].

Analyses were limited to white individuals for maximum statistical power to reduce potential for confounding due to the known population-based contribution to diversity in the HLA region. Participants were recruited and seen between 1984–2012. Written informed consent for genetic studies falling under the purview of the ADGC was obtained from all study participants, and institutional review boards at each site approved all aspects of this study. Specific consent for this study was obtained from the ADGC based on an application describing the proposed work.

ADNI. We also utilized data from 346 individuals recruited for participation in the ADNI study with data from SNP genotyping and longitudinal cognitive scores. All individuals included in this study had a minimum of two clinic assessments. At baseline, 120 individuals were cognitively normal (CN) older adults, 113 individuals were diagnosed with mild cognitive impairment (MCI), and 113 with AD. Of these, 163 individuals also had CSF measurements of plasma biomarkers available (S2 Table). The ADNI cohort is well characterized and has been used in previously published studies [27–29]. The clinical severity of symptoms in the MCI and AD groupings was measured using the Clinical Dementia Rating sum of boxes (CDR-SB) [30]. A clinician diagnosed each

participant using a structured protocol that utilized clinical judgment and neuropsychological tests that are provided in S1 Methods. The mean follow-up time was 3.15 $\pm$ 2.04 years for control participants (n = 91), 2.39 $\pm$ 1.71 years for participants with MCI (n = 148), and 1.37 $\pm$ 0.75 years for patients with AD (n = 69). Written informed consent was obtained from all study participants for research studies falling under the purview of ADNI, and the University of California, San Francisco Institutional Review Board approved all aspects of this study.

Genotype acquisition

UCSF MAC Cohort. Patient and control genotypes were obtained via genotyping on the Illumina Omni1-Quad array (Illumina, San Diego, California) using manufacturer's instructions. *APOE* genotype was determined with a TaqMan Allelic Discrimination Assay for the two SNPs, rs429358 and rs7412, on an ABI 7900HT Fast Real-Time PCR system (Applied Biosystems, Foster City, California) using the manufacturer's instructions.

ADGC. Details of genotyping in the 30 datasets that comprise the combined ADGC dataset are available online [26] and partially described in previously published papers [10].

ADNI. Haplotypes were determined using genotypes from the Human610-Quad BeadChip (Illumina, San Diego, California) as previously described [31]. *APOE* genotypes were determined by Cogenics (now Beckman Coulter; Pasadena, California).

CSF biomarker measurements

ADNI. Baseline CSF biomarkers levels were measured using the Human DiscoveryMAP panel developed by Rules Based Medicine (Myriad RBM; Austin,

Texas). The Human DiscoveryMAP panel is commercially available and measures a collection of metabolic, lipid, inflammatory, and other AD-relevant indicators. We limited our analyses to 28 immune proteins in the panel that were associated with inflammatory or immune processes (S1 List). The samples were processed and analyzed by Myriad RBM and checked for quality by the ADNI Biomarker core. CSF amyloid β 1–42 was measured using the AlzBio3 Luminex xMAP immunoassay (Innogenetics, Ghent, Belgium) according to previously described methods [32]. This method utilizes a monoclonal antibody specific for amyloid β 1–42 that is chemically bonded to color-coded beads along with analyte-specific detector antibodies. Additional details are available in S1 Methods.

Clinical assessments

ADNI. In this study, we analyzed two neuropsychological measures of cognitive function and one measure of clinical severity in ADNI participants. The Rey Auditory Verbal Learning Test (RAVLT) [33] is a test of verbal memory. It begins with the administrator reading a list of 15 unrelated words to the participant, who is then asked to verbally repeat as many of the words as they can. This happens for a total of five learning trials, and the administrator records the number of words correctly recalled after each trial. The test administrator then reads a set of 15 new words to the participant (interference word list), and, immediately following this, the participant is asked to recall as many of the first list of words as possible (immediate recall score). After a 30-min delay during which unrelated tests are administered, the participant is asked to recall as many words as possible from the initial list (delayed recall score). The RAVLT “forgetting score” is calculated as the difference between immediate recall versus delayed recall scores [33].

The forgetting score remains relatively stable over time in individuals with consistent memory function; the forgetting score tends to get smaller as the number of recalled items decreases. The 11-item Alzheimer's Disease Assessment Scale (ADAS) cognitive subscale assesses learning and memory, orientation, and several aspects of language including production, comprehension, and constructional and ideational praxis [34,35]. Higher scores indicate more impairment. Finally, the Clinical Dementia Rating (CDR) scale is a measure of three cognitive domains (memory, problem solving, and orientation) and three functional domains (self-care, community engagement, and hobbies). Information is collected directly from the study participant, as well as from a study informant. The scores for the six domains are combined into the CDR sum of boxes (CDR-SB) score [36].

Statistical analysis

Cohort demographic summary statistics. Summary statistics for participants' age, sex, age of onset, and *APOE-ε4* carrier status were calculated using R.

Imputation of HLA alleles. HLA genotypes were derived from chromosome 6 SNP data using an imputation program, HLA Genotype Imputation with Attribute Bagging (HIBAG) v1.3, which calculates predictions of genotype by averaging HLA-type posterior probabilities over an ensemble of classifiers built on bootstrap samples [37]. It relies on a training set of known HLA and SNP genotypes. We imputed the following HLA genes: *A*, *B*, *DRB1*, *DQA1*, and *DQB1*. For the UCSF MAC cohort and the ADGC merged dataset, a training set for four-digit resolution using ethnic-specific models for Europeans based on Omni1_Quad_v1_0_H was used. For the ADNI cohort clinical

biomarker analyses, we used four-digit resolution ethnic-specific models for Europeans derived from Illumina Human610-Quad v1.0.

Quality control of HLA imputation. Based on the distribution of posterior probabilities for each of the five imputed alleles (S1 Fig), we chose a call threshold (CT) of 0.75. As previous studies have shown that a CT of 0.5 leads to HIBAG prediction accuracies of 94.8%–99.2% for individuals of European ancestry [38], we expect our more stringent CT will correspond to similar or higher HIBAG prediction accuracies based on assumed accuracy of imputed ADGC SNPs. After excluding samples with any imputation probability below this cutoff at any locus, our final ADGC cohort size was 11,381.

Calculating locus and haplotype odds ratios (ORs). OR estimates for patients with AD and cognitively normal controls were calculated using a statistical package designed to specifically probe associations with the HLA (BIGDAWG), including tests of Hardy-Weinberg equilibrium and case–control association analyses for haplotypes as previously described [39]. Analyses were performed for each cohort (UCSF + ADGC) separately and in combination. As this was a fine-mapping study based on a previous genome-wide significant, and replicated, finding at *HLA-DRB5;* and considering that this study represents a first analysis of the highly polymorphic HLA region in the context of AD, a complex disease, we did not require a multiple testing correction. To strike a balance between reducing Type I error while also allowing for full exploration of the loci underlying this association with the MHC region, we implemented a stepwise assessment of HLA gene contributions to AD: using allelic information, we established a priori significance at $p < 0.05$ at the haplotype level based on the prior GWAS-significant results. We then examined the contingency table from which the haplotype result was

derived to identify the specific allele(s) contributing to the association signal. We accepted allele-level significance at $p < 0.05$ given the haplotype-level significance [40,41]. Based on a sample size of 11,690 in our combined UCSF+ADGC cohort, with 326 degrees of freedom and an alpha of 0.05, we had 64.1% power to detect an OR of 1.21 based on the haplotype frequencies of AD versus cognitively normal controls for the top associated five-allele haplotype.

Biomarker and cognitive data. Discrete and continuous demographic variables were compared across the ADNI cohort using chi-squared and ANOVA analyses, respectively. Linear mixed effects models were used to assess the relationship between the risk haplotype of interest and changes in the longitudinal cognitive measurements, ADAS and RAVLT, while controlling for baseline and time interactions of age, sex, education, baseline CDR-SB score (to account for baseline differences in clinical severity/diagnosis), and *APOE-ε4* carrier status. Use of linear mixed effects models allowed us to account for variable data missingness across participants by estimating subject-specific slopes. This enabled us to estimate cognitive changes for each individual despite varying numbers of visits. Missing data were omitted from the analyses, and all participants were required to have at least two time points to be included in the analysis. All interactions and main effects were modeled as fixed effects with random slopes and intercepts across individuals. The main effects of all variables were included in all longitudinal analyses but have been omitted from the definitions below to improve their clarity.

The linear mixed effects model for ADAS11 scores was defined as follows:

$\Delta\,\mathrm{ADAS} = \beta_0 + \beta_1 \Delta t + \beta_2 \mathrm{DR15} * \Delta t + \beta_3 \mathrm{Age} * \Delta t + \beta_4 \mathrm{Sex} * \Delta t + \beta_5 \mathrm{Education} * \Delta t + \beta_6 \mathrm{CDR\text{-}SB} * \Delta t + \beta_7 APOE\ \varepsilon 4 * \Delta t + e.$

The linear mixed effects model for the RAVLT forgetting score was defined as follows:

$\Delta\,\mathrm{RAVLT} = \beta_0 + \beta_1 \Delta t + \beta_2 \mathrm{DR15} * \Delta t + \beta_3 \mathrm{Age} * \Delta t + \beta_4 \mathrm{Sex} * \Delta t + \beta_5 \mathrm{Education} * \Delta t + \beta_6 \mathrm{CDR\text{-}SB} * \Delta t + \beta_7 APOE\ \varepsilon 4 * \Delta t + e.$

Cross-sectional CSF biomarker analyses. Linear models were used to test for an association between baseline CSF biomarker levels and the haplotype of interest. We controlled for age, sex, education, baseline CDR-SB score (to account for baseline differences in clinical severity/diagnosis), and *APOE-ε4* dosage.

Results

Five-allele haplotype analysis implicated DR15 in AD risk

The discovery UCSF cohort consisted of 309 individuals with clinically diagnosed AD and cognitively normal older adult controls (Table 1). Because of the small sample size, all imputed alleles were included in the haplotype analysis (HLA *A*, *B*, *DRB1*, *DQA1*, and *DQB1*). We performed association analysis on the four haplotypes with sufficient frequency in this small cohort. Of these four, one showed a significant association with AD risk: HLA *A*02:01~B*07:02~DRB1*15:01~DQA1*01:02~DQB1*06:02* (OR = 3.69; 95% confidence interval [CI] 1.16–13.69; $p = 0.01$) (Table 2).

After quality control, 11,381 individuals were available for analysis in the validation ADGC cohort (Table 1). Of the 318 haplotypes available for analysis, 12 five-allele haplotypes were significantly associated with AD risk ($p < 0.05$, Table 2). The

strongest association was HLA *A*03:01~B*07:02~DRB1*15:01~DQA1*01:02~ DQB1*06:02* (OR = 1.22 [1.08–1.38], $p = 8.5 \times 10^{-4}$). This haplotype differed from the UCSF finding by one allele, at *HLA-A*. The third most significant haplotype association in the ADGC cohort was *A*02:01~B*13:02~DRB1*07:01~DQA1*02:01~DQB1*02:02*, which showed a protective effect, (OR = 0.66 [0.50–0.89], $p = 4.2 \times 10^{-3}$). This haplotype shared the *HLA-A* allele associated with AD risk in the UCSF discovery analysis. The full *A*02:01~B*07:02~DRB1*15:01~DQA1*01:02~DQB1*06:02* haplotype associated with AD in the UCSF cohort was not significant in the ADGC cohort (p = 0.30).

Table 1. Cohort demographics.

Cohort	*n*	CN/AD	% Male	Age at onset
UCSF	309	191/118	46.3%	72.7 ± 9.0
ADGC	11,381	5,728/5,653	41.4%	74.0 ± 7.7

Mean ± standard deviation of age of onset indicates age of first reported symptoms. CN, cognitively normal; ADGC, Alzheimer's Disease Genetic Consortium merged dataset; UCSF, University of California, San Francisco Memory and Aging Center.

Table 2. Five-allele haplotype risk associations in UCSF and ADGC clinical cohorts.

Cohort Information	Haplotype: *A~B~DRB1~DQA1~DQB1*	OR (95% CI)	*p*-Value	CN frequency	AD frequency
UCSF	01:01~08:01~03:01~05:01~02:01	0.87 (0.45–1.64)	0.66	0.0864	0.0763
n = 309	**02:01~07:02~15:01~01:02~06:02**	**3.69 (1.16–13.69)**	**0.01**	**0.0131**	**0.0466**
191 CN; 118 AD	02:01~08:01~03:01~05:01~02:01	0.64 (0.15–2.26)	0.45	0.0262	0.0170
4 haplotypes analyzed	03:01~07:02~15:01~01:02~06:02	0.72 (0.28–1.68)	0.42	0.0524	0.0381
ADGC	01:01~08:01~07:01~02:01~03:03	0.43 (0.16–1.02)	0.04	0.0017	0.0007
	02:01~07:02~15:01~01:02~06:02	**1.08 (0.93–1.25)**	**0.30**	**0.0320**	**0.0345**
n = 11,381	**02:01~13:02~07:01~02:01~02:02**	**0.66 (0.50–0.89)**	**4.2×10^{-3}**	**0.0107**	**0.0072**
	02:01~15:01~07:01~02:01~02:02	0.39 (0.14–0.99)	0.03	0.0016	0.0006
5,728 CN; 5,653 AD	02:01~44:02~13:01~01:03~06:03	1.44 (1.03–2.03)	0.03	0.0054	0.0078
	02:01~57:01~07:01~02:01~03:03	1.31 (1.01–1.69)	0.04	0.0095	0.0124
318 haplotypes analyzed	03:01~07:02~12:01~05:05~03:01	0.30 (0.09–0.84)	0.01	0.0015	0.0004
	03:01~07:02~15:01~01:02~06:02	**1.22 (1.08–1.38)**	**8.5×10^{-4}**	**0.0472**	**0.0570**
	11:01~35:01~07:01~02:01~02:02	0.31 (0.07–1.01)	0.03	0.0011	0.0004
	24:02~38:01~13:01~01:03~06:03	**0.14 (0.02–0.63)**	**2.9×10^{-3}**	**0.0012**	**0.0002**
	24:02~44:05~01:01~01:01~05:01	4.56 (0.94–43.38)	0.03	0.0002	0.0008
	29:02~58:01~08:04~04:01~04:02	4.56 (0.94–43.38)	0.03	0.0002	0.0008
	68:01~44:02~01:01~01:01~05:01	1.96 (0.99–4.04)	0.04	0.0012	0.0024

All analyzed haplotype association results (regardless of significance) are reported for the UCSF cohort, with the significant ($p < 0.05$) finding in bold. Nonsignificant results ($p > 0.05$) are shown in italics. For the ADGC cohort, all significant ($p < 0.05$) haplotype association results are reported, in addition to the results for the single significant haplotype from the UCSF cohort (not significant in the ADGC analysis [$p = 0.30$], in italics). The top three most significant ADGC and findings are highlighted in blue. In addition to odds ratio (OR) with 95% confidence interval (CI), a breakdown of the haplotype frequency in individuals with Alzheimer disease (AD) versus cognitively normal (CN) older adult controls is also provided.

In combined analysis of both the UCSF and ADGC cohorts, 326 haplotypes were available for analysis (additional haplotypes beyond the 4 + 318 haplotypes analyzed in the separate UCSF and ADGC cohorts resulted when sufficient numbers of AD and CN controls for rare haplotypes became available in the combined UCSF + ADGC dataset). HLA *A*03:01~B*07:02~DRB1*15:01~DQA1*01:02~DQB1*06:02* (OR = 1.21 [1.08–1.37), $p = 9.6 \times 10^{-4}$) and *A*02:01~B*13:02~DRB1*07:01~DQA1*02:01~DQB1*02:02*, [OR = 0.66 [0.49–0.88], $p = 3.8 \times 10^{-3}$] remained as two of the three most significant associations with AD (Table 3). Locus-level analyses of the combined cohort showed independent AD associations of *B*07:02*, *DRB1*15:01*, *DQA1*01:02*, and *DQB1*06:02* (Table 4).

Table 3. Five-allele haplotype risk associations in combined UCSF and dataset.

Cohort Information	Haplotype: *A~B~DRB1~DQA1~DQB1*	OR (95% CI)	*p*-Value	CN frequency	AD frequency
ADGC + UCSF	01:01~08:01~07:01~02:01~03:03	0.43 (0.16–1.03)	0.04	0.0016	0.0007
	02:01~07:02~01:01~01:01~05:01	1.78 (0.98–3.32)	0.04	0.0016	0.0029
	02:01~07:02~15:01~01:02~06:02	*1.08 (0.94–1.25)*	*0.28*	*0.0317*	*0.0342*
n = 11,690	**02:01~13:02~07:01~02:01~02:02**	**0.66 (0.49–0.88)**	**3.8×10^{-3}**	**0.0106**	**0.0070**
	02:01~15:01~07:01~02:01~02:02	0.38 (0.13–0.94)	0.02	0.0016	0.0006
5,919 CN; 5,717 AD	02:01~18:01~07:01~02:01~02:02	3.42 (0.88–19.35)	0.047	0.0003	0.0009
	02:01~44:02~13:01~01:03~06:03	1.49 (1.07–2.10)	0.01	0.0052	0.0078
326 haplotypes analyzed	02:01~57:01~07:01~02:01~03:03	1.32 (1.02–1.72)	0.03	0.0092	0.0121
	03:01~07:02~12:01~05:05~03:01	0.32 (0.09–0.92)	0.02	0.0014	0.0004
	03:01~07:02~15:01~01:02~06:02	**1.21 (1.08–1.37)**	**9.6×10^{-4}**	**0.0476**	**0.0573**
	11:01~35:01~07:01~02:01~02:02	0.29 (0.07–0.93)	0.02	0.0012	0.0003
	24:02~38:01~13:01~01:03~06:03	**0.15 (0.02–0.64)**	**3.2×10^{-3}**	**0.0012**	**0.0002**
	24:02~44:05~01:01~01:01~05:01	5.13 (1.09–48.17)	0.02	0.0002	0.0009
	29:02~58:01~08:04~04:01~04:02	4.62 (0.96–43.93)	0.03	0.0002	0.0008
	68:01~44:02~01:01~01:01~05:01	1.91 (0.96–3.95)	0.048	0.0012	0.0023

= All significant ($p < 0.05$) haplotype association results from the combined UCSF + ADGC dataset are reported, in addition to the results for the single significant haplotype from the UCSF cohort (not significant in the the ADGC+UCSF analysis [$p = 0.28$] in italics). The top three most significant ADGC+UCSF findings are highlighted in blue. In addition to odds ratio (OR) with 95% confidence interval (CI), a breakdown of the haplotype frequency in individuals with Alzheimer disease (AD) versus cognitively normal (CN) older adult controls is also provided.

Table 4. Individual alleles with significant risk associations in combined cohort.

Class I Loci	OR (95% CI)	*p*-Value	CN frequency	AD frequency	European population frequency estimate
A*23:01	0.81 (0.67–0.98)	0.03	0.0223	0.0182	0.0168
A*33:03	1.97 (0.97–4.21)	0.04	0.0011	0.0022	0.0013
B*07:02	**1.07 (1.00–1.15)**	**0.04**	**0.1629**	**0.1727**	**0.1400**
B*15:01	0.87 (0.78–0.98)	0.02	0.0627	0.0553	0.0665
B*41:01	2.15 (1.00–4.88)	0.03	0.0009	0.0020	0.0038
B*57:01	1.15 (1.01–1.30)	0.03	0.0406	0.0462	0.0383
Class II Loci	**OR (95% CI)**	***p*-Value**	**CN frequency**	**AD frequency**	**European population frequency estimate**
DRB1*15:01	**1.08 (1.01–1.15)**	**0.03**	**0.1795**	**0.1907**	**0.1444**
DQA1*01:02	**1.06 (1.00–1.13)**	**0.04**	**0.2373**	**0.2487**	**not available**
DQB1*06:02	**1.08 (1.01–1.15)**	**0.03**	**0.1782**	**0.1895**	**0.1425**

All significant loci results ($p < 0.05$) for combined UCSF and ADGC cohort (n = 11,690). Alleles present in one of the top three most significant ($p < 0.01$) five-allele haplotypes (Table 3) are shown in bold. The number of analyzed alleles differed by loci (A: n = 23, B: n = 39, DQB1: n = 27, DQA1: n = 13, DQB1: n = 15). In addition to OR with 95% CI, a breakdown of allele frequency in individuals with Alzheimer's disease (AD) versus cognitively normal (CN) older adult controls is also provided in addition to the expected frequency in populations of European descent [42].

Five-allele haplotype contributed to AD risk independently of APOE-ε4 and may be driven by ε4-negative individuals

We next assessed whether the strong genetic AD risk factor *APOE-ε4* can account for the most significant five-allele haplotype association we identified in the combined UCSF+ADGC cohort. We recoded individuals as carriers or non-carriers of the *A*03:01~B*07:02~DRB1*15:01~DQA1*01:02~DQB1*06:02* haplotype and assessed in a logistic regression framework the independent contributions of the risk haplotype, *APOE-ε4* carrier status, and whether there was an interaction between the two. As expected, *APOE-ε4* was strongly associated with AD risk ($p < 2 \times 10^{-16}$). The five-allele risk haplotype remained a significant contributor to AD risk ($p = 0.036$), but there was no statistically significant interaction between haplotype and *APOE-ε4* ($p = 0.19$). However, dividing the cohort by *APOE-ε4* carrier status showed that the frequency of the *A*03:01~B*07:02~DRB1*15:01~DQA1*01:02~DQB1*06:02* haplotype was higher only in individuals with AD who are negative for ε4 (Table 5). Analysis of variance performed separately in ε4 carriers and non-carriers resulted in a significant association of the five-allele haplotype with AD only in ε4-negative individuals ($p = 0.036$ in ε4 non-carriers; $p = 0.90$ in ε4 carriers).

Table 5. AD/CN control distribution by *APOE-ε4* and *A*03:01~B*07:02~DRB1*15:01~DQA1*01:02~DQB1*06:02* haplotype carrier status.

Haplotype status	*APOE-ε4* +		*APOE-ε4* -	
	AD	**CN**	**AD**	**CN**
*A*03:01~B*07:02~DRB1*15:01~DQA1*01:02~DQB1*06:02* carriers	0.726	0.274	0.391	0.609
*A*03:01~B*07:02~DRB1*15:01~DQA1*01:02~DQB1*06:02* non-carriers	0.723	0.277	0.344	0.656
p-Value (ANOVA)	0.901		**0.036**	

Alzheimer's disease (AD)/cognitively normal (CN) distribution by *APOE-ε4* status and risk haplotype carrier status. Of the full cohort, 9,517 individuals had information on *APOE* genotype and were included in this analysis. Individuals with either one or two ε4 alleles were classified as *APOE-ε4* positive.

Separate class I and class II haplotype analyses corroborated A*03:01~B*07:02 and DR15 in AD risk

Given the different roles of HLA receptors in recognizing endogenous (class I) or exogenous (class II) ligands, we also assessed class I (*HLA A~B*) and class II (*HLA DRB1~DQA1~DQB1*) haplotypes separately for their role in AD risk in the combined UCSF+ADGC cohort. Of 202 analyzed class I haplotypes, ten two-allele haplotypes were significantly associated with AD ($p < 0.05$), including *A*03:01~B*07:02* ($p = 0.03$, OR = 1.1 [1.0–1.2]) (Table 6). Only one three-allele class II haplotype (out of 30 analyzed) was associated with AD risk, *DR15* ($p = 0.025$, OR = 1.1 [1.0–1.2]). Together, these two separate haplotypes represent the most strongly associated five-allele haplotype identified in the combined analysis.

Table 6. Separate class I and class II haplotypes with significant risk associations in combined cohort.

Class I Haplotypes *A~B*	OR (95% CI)	*p*-Value	CN frequency	AD frequency
01:01~57:01	1.21 (1.01–1.46)	0.04	0.0187	0.0225
02:01~13:02	**0.66 (0.51–0.86)**	**1.4×10^{-3}**	**0.0130**	**0.0087**
03:01~07:02	**1.11 (1.01–1.23)**	**0.03**	**0.0703**	**0.0777**
03:01~15:01	0.63 (0.45–0.87)	4.1×10^{-3}	0.0083	0.0052
11:01~15:01	0.54 (0.29–0.99)	0.03	0.0029	0.0016
24:02~38:01	**0.36 (0.13–0.88)**	**0.01**	**0.0017**	**0.0006**
26:01~39:01	0.07 (0–0.48)	9.4×10^{-4}	0.0012	0.0001
26:01~44:02	0.15 (0.02–0.64)	3.2×10^{-3}	0.0012	0.0002
32:01~14:02	4.62 (0.96–43.93)	0.03	0.0002	0.0008
68:01~40:01	0.48 (0.30–0.75)	7.0×10^{-4}	0.0054	0.0026
Class II Haplotypes *DRB1~DQB1~DQA1*	**OR (95% CI)**	***p*-Value**	**CN frequency**	**AD frequency**
15:01~01:02~06:02	**1.08 (1.01–1.15)**	**0.03**	**0.1781**	**0.1894**

All significant ($p < 0.05$) class I (*A~B*) and class II (*DRB1~DQA1~DQB1*) haplotypes for combined UCSF and ADGC cohorts (n = 11,690). Class I and class II haplotypes present in one of the top three most significant ($p < 0.01$) five-allele haplotypes (Table 2) are shown in bold. In total, 202 class I haplotypes and 30 class II haplotypes were analyzed. In addition to OR with 95% CI, a breakdown of allele frequency in individuals with Alzheimer's disease (AD) versus cognitively normal (CN) older adult controls is also provided.

*Class I haplotype A*03:01~B*07:02 was associated with baseline CSF amyloid levels*

We utilized a subset of the ADNI cohort with genetic and cognitive data available to assess the disease-specific relevance of the class I *A*03:01~B*07:02* and class II *DR15* haplotypes across the AD spectrum, including cognitively normal controls, individuals with MCI, and those with AD. We analyzed the two haplotypes separately to assess whether class I and class II risk-associated haplotypes were correlated with similar or different clinical measures of AD. The cohort was balanced with respect to age, sex, and haplotype distributions (Table 7). The cohort was significantly different with respect to education and number of time points and showed expected differences in CDR-SB baseline score, *APOE-ε4* carrier status, ADAS baseline score, and RAVLT forgetting baseline score.

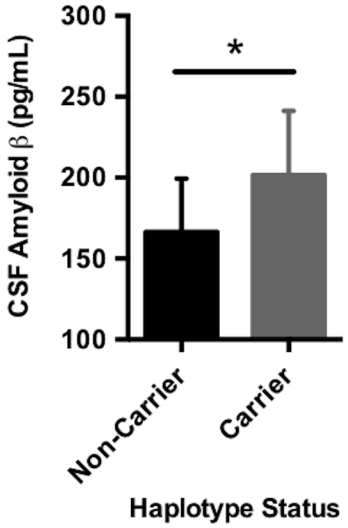
Table 7. Summary statistics for ADNI participants with longitudinal cognitive measures.

	CN	MCI	AD	*p*-Value
n	120	113	113	
Age (years)	75.6 ± 4.87	74.0 ± 6.17	75.9 ± 6.72	NS
Sex (% female)	45.8%	35.4%	47.8%	NS
Education (years)	16.0 ± 2.7	15.5 ± 2.9	14.6 ± 3.0	<0.001
CDR-SB score	0.02 ± 0.1	1.7 ± 0.9	4.3 ± 1.7	<0.001
APOE-ε4 carrier (%)	31.7%	67.3%	69.0%	<0.001
Haplotype dose (number of single / number of double)	33/0	30/3	33/3	NS
Time points	7.4 ± 2.8	6.6 ± 2.4	3.8 ± 0.8	<0.001
ADAS score (baseline)	6.0 ± 2.9	12.2 ± 4.1	18.5 ± 6.1	<0.001
RAVLT forgetting score (baseline)	3.5 ± 2.8	4.8 ± 2.3	4.5 ± 1.9	<0.001

Descriptive data are summarized by diagnostic category. Values represent the mean ± standard deviation and the percent or number of participants in a given diagnostic category. Two-tailed p-values were from analysis of variance (continuous traits) or chi-square (categorical values) tests by diagnostic group. NS, not significant (p > 0.05)

Carrying *A*03:01~B*07:02* was associated with higher baseline levels of amyloid β as measured in CSF (Fig 1, $p = 0.01$). Traditionally, CSF amyloid levels are inversely correlated with amyloid burden in the brain; our results suggest that carrying *A*03:01~B*07:02* is correlated with lower amyloid levels in the brain [43]. This is observed despite the fact that there were no statistically significant differences in baseline clinical or biomarker measures in patients with versus without the risk haplotype (S3 Table, S2 Fig). *A*03:01~B*07:02* was not associated with any other baseline measures and was not associated with change in longitudinal measures over time.

Figure 1. Carrying the *A*03:01~B*07:02* risk haplotype was associated with CSF (cerebrospinal fluid) amyloid β.



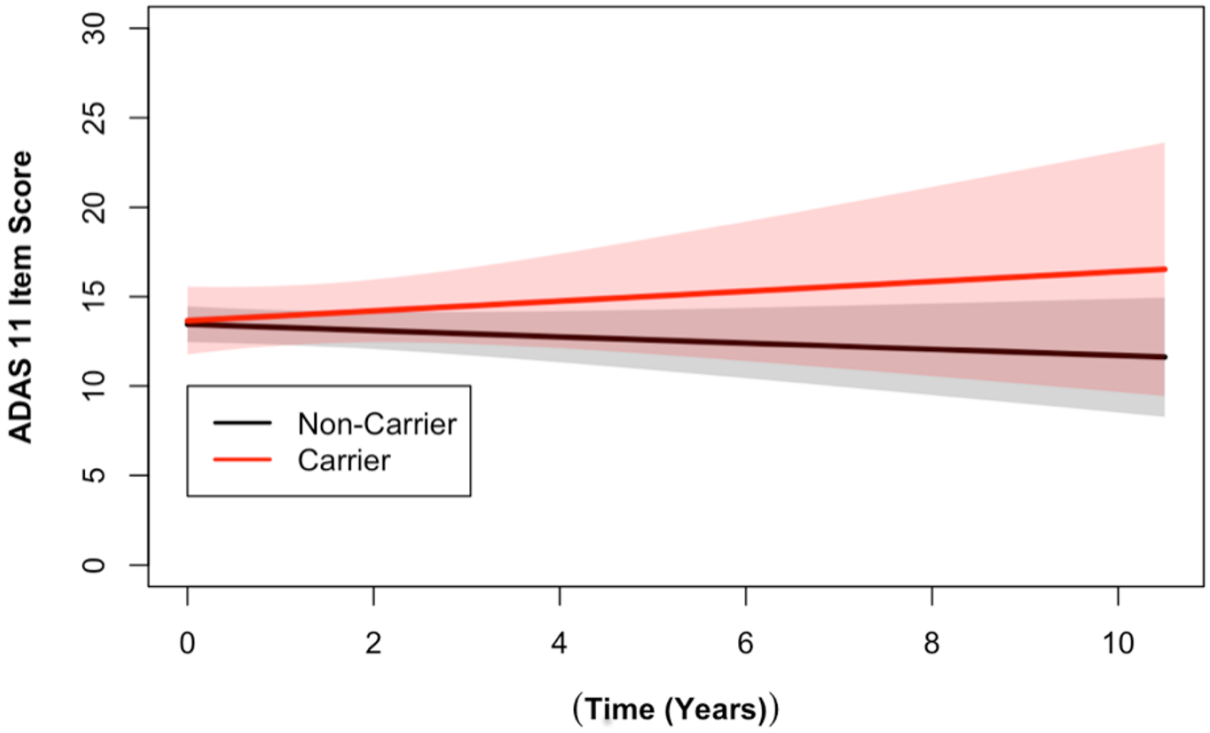
CSF amyloid β levels were on average higher in carriers of the A*03:01~B*07:02 haplotype, suggesting that haplotype carriers may have lower average intracranial amyloid pathological burden compared to noncarriers. The plotted points are best linear unbiased predictions from a multiple regression model, which controlled for age, sex, education, CDR-SB score, and *APOE-ε4* status. Data shown are the mean ± standard deviation (SD).

DR15 risk haplotype correlated with worse cognitive decline and greater baseline inflammation across the AD spectrum

Longitudinal analysis of cognitive data identified a statistically significant association between the number of alleles of the *DR15* risk haplotype and ADAS cognitive scores ($p = 0.03$), as well as with RAVLT forgetting scores ($p = 0.02$, S4 Table). The *DR15* haplotype was associated with worse decline over time on both measures, corresponding to increasing longitudinal ADAS cognitive scores and decreasing longitudinal RAVLT forgetting scores over time (shown relative to non-carriers in Figs 2 and 3). *DR15*-associated changes in cognitive trajectory occurred despite the fact that there were no baseline differences in clinical severity or cognitive function in patients with AD based on *DR15* carrier status (S5 Table). In addition, baseline biomarker measures most relevant to AD were similar in both patients with AD who are *DR15* carriers and those who are noncarriers (S5 Table, S3 Fig), indicating that all patients had equivalent baseline disease severity.

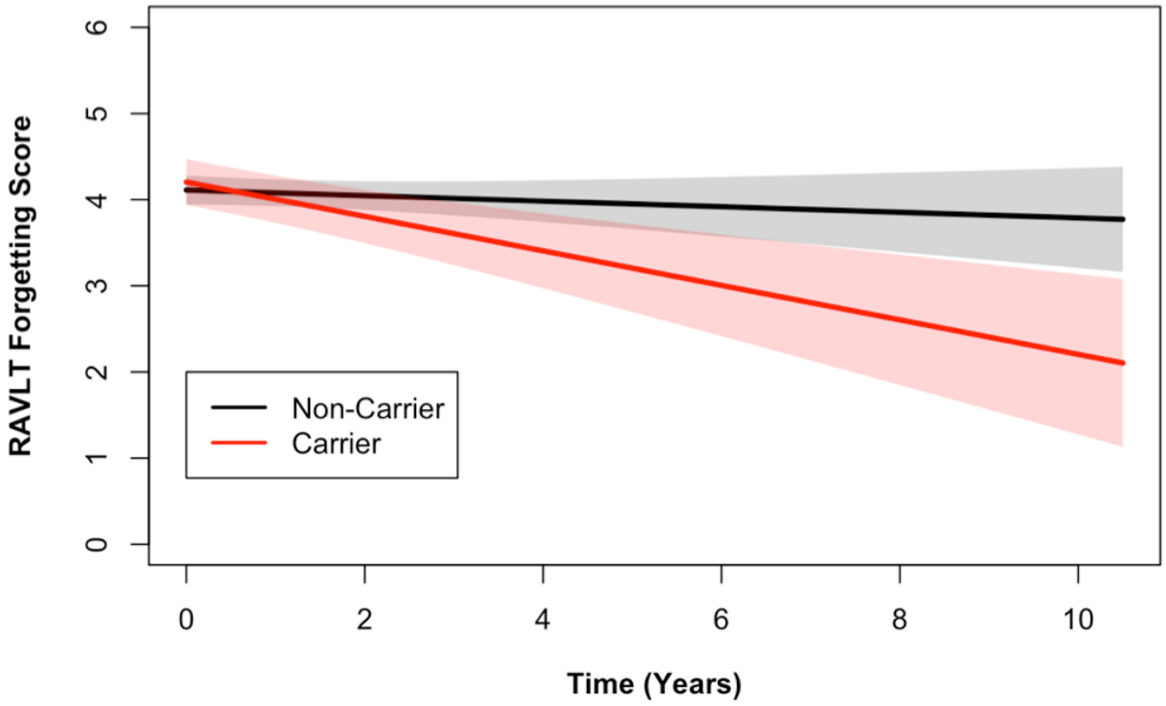
In a subset of individuals who also had baseline CSF data available (S2 Table), we tested whether the *DR15* risk haplotype altered any biomarker measures of immunological function and inflammation. We tested 28 analyte levels related to immune function and inflammation (S1 List). At baseline, there was an association between chemokine CC-4 (CC4) and age ($p = 0.02$, S4 Fig), as well as CC4 with dose of *DR15* risk haplotype ($p = 5.18 \times 10^{-3}$, Fig 4, S6 Table). Although not reaching strict statistical significance after adjustment for the 28 biomarkers tested (at Bonferroni adjusted $p < 1.79 \times 10^{-3}$), this analysis provides suggestive biomarker evidence of heightened baseline inflammation in individuals carrying the *DR15* risk haplotype.

Figure 2. *DR15* haplotype carriers showed greater change over time on the ADAS cognitive assessment when compared to noncarriers.



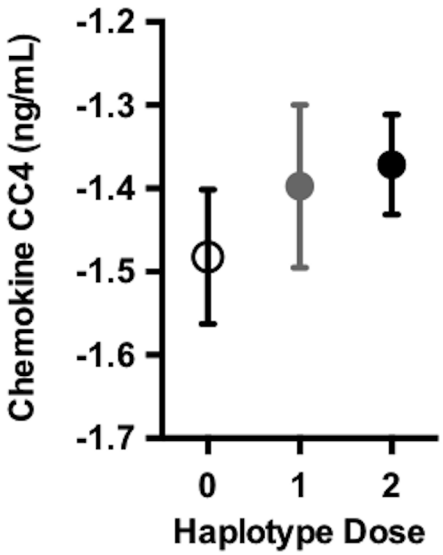
Longitudinal ADAS 11-item cognitive subscale scores from the ADNI cohort are shown. The ADAS broadly measures cognitive functions impaired in AD [34], with higher scores representing more cognitive impairment. DR15 haplotype carriers (in red) showed worse cognitive function over time when compared to noncarriers (in black) ($p = 0.03$). The plotted data represent the best linear unbiased prediction results from the regression model specified (see Methods) with 95% CIs (shaded regions).

Figure 3. *DR15* haplotype carriers declined more on the RAVLT forgetting score when compared to non-carriers.



Longitudinal RAVLT measurements from the ADNI cohort are shown. The RAVLT forgetting score is defined as the difference between the delayed recall and immediate recall scores on the RAVLT and represents a measure of memory consolidation. Over time, DR15 risk haplotype carriers showed more change on the forgetting score (i.e., more forgetting) than non-carriers. The plotted data represent the best linear unbiased prediction results from the regression model specified (see Methods) with 95% CIs (shaded regions).

Figure 4. *DR15* dosage was associated with higher baseline levels of chemokine CC4.



As the number of DR15 risk haplotype alleles increases, there were higher average levels of chemokine CC4, suggesting higher levels of inflammation at baseline. Chemokine CC4 levels are quality controlled and transformed as described in S1 Methods. The plotted points are partial residuals with 95% confidence bands provided in shading.

HLA haplotype risk effects differed by sex

Given previous reports of greater risk effects of *DR15* in female patients with multiple sclerosis (MS) [44] and the stronger effect of *APOE-ε4* in females [45], we assessed whether men versus women showed similar or different HLA haplotype associations with AD risk. When split by sex, two of the three most significant five-allele haplotypes from the combined sex analysis were significant in an individual sex. The five-allele haplotype *A*03:01~B*07:02~DRB1*15:01~DQA1*01:02~DQB1*06:02* was significant only in men (OR 1.31 [1.09–1.58], $p = 0.0035$) (Table 8). However, *A*02:01~B*13:02~DRB1*07:01~DQA1*02:01~DQB1*02:02* was significant only in women (OR 0.68 [0.46–0.99], $p = 0.034$) (Table 8). Similar findings appeared in separate class I and class II haplotype analyses. Only men showed significant associations with class I haplotype *A*03:01~B*07:02* ($p = 0.027$), and only women showed significant associations with *A*02:01~B*13:02* ($p = 0.0049$) (Table 9). Finally, class II haplotype *DR15* was only significantly associated with AD risk in men ($p = 0.01$) (Table 9). Locus-level analyses were consistent, with only men showing significant associations with ten alleles, including *B*07:02* ($p = 0.013$), *DRB1*15:01* ($p = 0.0096$), *DQA1*01:02* ($p = 0.029$), and *DQB1*06:02* ($p = 0.01$) (Table 10). There were four individual alleles associated with AD risk in women, none of which were components of any of the top three significant five-allele haplotypes in the combined sex analysis (Table 10).

Table 8. Five-allele haplotypes with significant risk associations in individual sexes.

Significant Haplotypes *A~B~DRB1~DQA1~DQB1*	Men n = 4860 (2369 cases, 2491 controls) OR (95% CI)	*p*-Value	Frequency CN	AD	Women n = 6930 (3402 cases, 3428 controls) OR (95% CI)	*p*-Value	Frequency CN	AD
01:01~37:01~15:01~01:02~06:02	**0.33 (0.11 - 0.86)**	**0.01**	**0.004**	**0.001**	1.27 (0.67 - 2.46)	0.43	0.003	0.004
02:01~07:02~01:01~01:01~05:01	1.42 (0.73 - 2.84)	0.27	0.003	0.005	**2.52 (1.06 - 6.63)**	**0.02**	**0.001**	**0.003**
02:01~13:02~07:01~02:01~02:02	**0.68 (0.42 - 1.09)**	**0.09**	**0.010**	**0.007**	**0.68 (0.46 - 0.99)**	**0.03**	**0.011**	**0.007**
02:01~15:01~07:01~02:01~02:02	n/a	n/a	0.001	0.001	**0.25 (0.05 - 0.93)**	**0.02**	**0.002**	**0.000**
02:01~27:05~08:01~04:01~04:02	1.05 (0.37 - 2.99)	0.92	0.002	0.002	**0.3 (0.09 - 0.84)**	**0.01**	**0.002**	**0.001**
02:01~40:01~04:01~03:01~03:02	0.53 (0.12 - 1.96)	0.29	0.002	0.001	**2.74 (1.10 - 7.72)**	**0.02**	**0.001**	**0.003**
02:01~44:02~03:01~05:01~02:01	**5.79 (1.26 - 53.79)**	**0.01**	**0.000**	**0.002**	1.16 (0.52 - 2.66)	0.69	0.002	0.002
02:01~44:02~13:01~01:03~06:03	**1.71 (1.00 - 2.99)**	**0.04**	**0.005**	**0.008**	1.38 (0.89 - 2.16)	0.13	0.006	0.008
02:01~52:01~15:02~01:03~06:01	n/a	n/a	0.000	0.000	**5.04 (1.07 - 47.36)**	**0.02**	**0.000**	**0.001**
02:01~57:01~07:01~02:01~03:03	**1.81 (1.17 - 2.84)**	**$4.74*10^{-3}$**	**0.007**	**0.013**	1.10 (0.08 - 1.52)	0.55	0.011	0.012
03:01~07:02~15:01~01:02~06:02	**1.31 (1.09 - 1.58)**	**$3.46*10^{-3}$**	**0.045**	**0.058**	**1.12 (0.96 - 1.30)**	**0.16**	**0.049**	**0.055**
03:01~15:01~13:01~01:03~06:03	**0.28 (0.07 - 0.88)**	**0.02**	**0.003**	**0.001**	1.21 (0.48 - 3.13)	0.66	0.001	0.002
03:01~27:05~01:01~01:01~05:01	n/a	n/a	0.000	0.001	**0.07 (0 - 0.47)**	**$8.27*10^{-4}$**	**0.002**	**0.000**
03:01~57:01~07:01~02:01~03:03	**0.42 (0.16 - 1.00)**	**0.03**	**0.004**	**0.002**	1.73 (0.63 - 5.19)	0.24	0.001	0.002
11:01~35:01~07:01~02:01~02:02	n/a	n/a	0.000	0.000	**0.15 (0.02 - 0.68)**	**$4.70*10^{-3}$**	**0.002**	**0.000**
11:01~51:01~01:01~01:01~05:01	n/a	n/a	0.000	0.000	**0.22 (0.02 - 1.08)**	**0.04**	**0.001**	**0.000**
26:01~27:05~01:01~01:01~05:01	**3.42 (1.06 - 14.42)**	**0.02**	**0.001**	**0.003**	0.76 (0.28 - 1.95)	0.52	0.002	0.001
68:01~44:02~11:01~05:05~03:01	**0.42 (0.16 - 1.00)**	**0.03**	**0.004**	**0.002**	1.76 (0.69 - 4.86)	0.19	0.001	0.002

All significant 5-allele haplotype results (*p<0.05) for combined UCSF + ADGC cohort (n = 11,690) when males and females are analyzed separately. Two of the three most significant 5-allele haplotypes from the combined analysis (males + females) were significant in an individual sex analysis and are highlighted in this table in blue. Non-significant results are shown in grey and significant findings are shown in bold. In addition to odds ratio (OR) with 95% confidence interval (CI), a breakdown of haplotype frequency in Alzheimer's disease cases versus healthy older adult controls is also provided.

Table 9. Class I and class II haplotypes with significant risk associations in individual sexes.

Class I *A~B*	Males n = 4860 (2369 cases, 2491 controls) OR (95% CI)	*p*-Value	Frequency CN	AD	Females n = 6930 (3402 cases, 3428 controls) OR (95% CI)	*p*-Value	Frequency CN	AD
02:01~13:02	0.67 (0.43 - 1.03)	0.05	0.012	0.008	**0.63 (0.45 - 0.88)**	**$4.89*10^{-3}$**	**0.014**	**0.009**
02:01~18:01	1.23 (0.77 - 2.00)	0.36	0.007	0.009	**1.53 (1.03 - 2.28)**	**0.03**	**0.007**	**0.010**
02:01~55:01	0.53 (0.14 - 1.69)	0.23	0.002	0.001	**5.04 (1.07 - 47.36)**	**0.02**	**0.000**	**0.001**
03:01~07:02	**1.19 (1.02 - 1.39)**	**0.03**	**0.069**	**0.080**	1.06 (0.93 - 1.20)	0.40	0.072	0.075
03:01~15:01	**0.44 (0.25 - 0.75)**	**$1.18*10^{-3}$**	**0.010**	**0.004**	0.88 (0.56 - 1.37)	0.54	0.007	0.006
03:01~27:05	0.89 (0.44 - 1.80)	0.73	0.004	0.004	**0.50 (0.24 - 1.02)**	**0.04**	**0.004**	**0.002**
03:01~40:01	1.90 (0.83 - 4.60)	0.10	0.002	0.004	**0.43 (0.20 - 0.88)**	**0.01**	**0.004**	**0.002**
03:01~55:01	1.40 (0.43 - 4.91)	0.53	0.001	0.002	**4.54 (0.94 - 43.15)**	**0.03**	**0.000**	**0.001**
11:01~15:01	1.05 (0.28 - 3.94)	0.93	0.001	0.001	**0.44 (0.20 - 0.93)**	**0.02**	**0.004**	**0.002**
11:01~44:02	0.83 (0.43 - 1.59)	0.55	0.005	0.004	**0.46 (0.20 - 0.98)**	**0.03**	**0.004**	**0.002**
11:01~44:03	n/a	n/a	0.001	0.001	**2.69 (1.00 - 8.40)**	**0.03**	**0.001**	**0.002**
24:02~38:01	**0.10 (0 - 0.66)**	**0.01**	**0.002**	**0.000**	0.67 (0.20 - 2.11)	0.45	0.001	0.001
68:01~40:01	**0.43 (0.17 - 0.97)**	**0.03**	**0.004**	**0.002**	**0.48 (0.26 - 0.86)**	**0.01**	**0.006**	**0.003**
Class II *DRB1~DQB1~DQA1*	**OR (95% CI)**	***p*-Value**	**Frequency CN**	**AD**	**OR (95% CI)**	***p*-Value**	**Frequency CN**	**AD**
10:01~01:05~05:01	**2.21 (1.00 - 5.27)**	**0.03**	**0.002**	**0.004**	0.88 (0.47 - 1.65)	0.67	0.004	0.003
11:01~05:05~03:01	**0.66 (0.46 - 0.94)**	**0.02**	**0.017**	**0.011**	1.01 (0.75 - 1.35)	0.96	0.014	0.014
12:01~05:05~03:01	1.05 (0.77 - 1.45)	0.74	0.017	0.018	**0.72 (0.55 - 0.94)**	**0.01**	**0.020**	**0.015**
13:02~01:02~06:09	1.03 (0.70 - 1.53)	0.86	0.011	0.012	**0.69 (0.47 - 0.99)**	**0.04**	**0.011**	**0.008**
15:01~01:02~06:02	**1.14 (1.03 - 1.27)**	**0.01**	**0.173**	**0.193**	1.04 (0.95 - 1.13)	0.43	0.182	0.187

All significant class I (2-allele) and class II (3-allele) haplotype results (*p<0.05) for combined UCSF + ADGC cohort (n = 11,690) when males and females are analyzed separately. Class I and class II haplotypes present in one of the three most significant 5-allele from the combined analysis (males + females) are highlighted in this table in blue. Non-significant results are shown in grey and significant associations are shown in bold. In addition to odds ratio (OR) with 95% confidence interval (CI), a breakdown of haplotype frequency in Alzheimer's disease cases versus healthy older adult controls is also provided.

Table 10. Individual alleles with significant risk associations in individual sexes.

	Males n = 4860 (2369 cases, 2491 controls)				Females n = 6930 (3402 cases, 3428 controls)			
			Frequency				Frequency	
Class I Loci	**OR (95% CI)**	***p*-Value**	**CN**	**AD**	**OR (95% CI)**	***p*-Value**	**CN**	**AD**
A*23:01	0.91 (0.68 - 1.22)	0.51	0.0209	0.0190	**0.75 (0.59 - 0.96)**	**0.02**	**0.0233**	**0.0176**
A*32:01	**1.32 (1.05 - 1.67)**	**0.02**	**0.0279**	**0.0365**	0.96 (0.80 - 1.16)	0.70	0.0349	0.0337
B*07:02	**1.14 (1.03 - 1.27)**	**0.01**	**0.1626**	**0.1815**	1.03 (0.94 - 1.12)	0.59	0.1631	0.1665
B*15:01	**0.83 (0.70 - 0.99)**	**0.03**	**0.0634**	**0.0532**	0.91 (0.79 - 1.05)	0.18	0.0621	0.0567
B*39:01	1.09 (0.64 - 1.86)	0.75	0.0060	0.0065	**0.58 (0.36 - 0.91)**	**0.01**	**0.0079**	**0.0046**
B*58:01	**2.11 (1.12 - 4.12)**	**0.01**	**0.0032**	**0.0068**	0.78 (0.45 - 1.32)	0.32	0.0051	0.0040
			Frequency				Frequency	
Class II Loci	**OR (95% CI)**	***p*-Value**	**CN**	**AD**	**OR (95% CI)**	***p*-Value**	**CN**	**AD**
DRB1*10:01	**2.21 (1.00 - 5.27)**	**0.03**	**0.0020**	**0.0044**	0.88 (0.47 - 1.65)	0.67	0.0035	0.0031
DRB1*11:01	**0.66 (0.46 - 0.94)**	**0.02**	**0.0173**	**0.0114**	1.01 (0.75 - 1.35)	0.96	0.0142	0.0143
DRB1*12:01	1.05 (0.77 - 1.45)	0.74	0.0169	0.0177	**0.72 (0.55 - 0.94)**	**0.01**	**0.0203**	**0.0147**
DRB1*15:01	**1.15 (1.03 - 1.27)**	**$9.60*10^{-3}$**	**0.1744**	**0.1948**	1.03 (0.94 - 1.13)	0.49	0.1832	0.1878
DQA1*01:02	**1.11 (1.01 - 1.22)**	**0.03**	**0.2355**	**0.2545**	1.03 (0.96 - 1.12)	0.41	0.2386	0.2447
DQA1*01:05	**2.21 (1.00 - 5.27)**	**0.03**	**0.0020**	**0.0044**	0.88 (0.47 - 1.65)	0.67	0.0035	0.0031
DQB1*06:02	**1.14 (1.03 - 1.27)**	**0.01**	**0.1728**	**0.1929**	1.03 (0.95 - 1.13)	0.45	0.1820	0.1871
DQB1*06:09	1.03 (0.70 - 1.53)	0.86	0.0112	0.0116	**0.69 (0.47 - 0.99)**	**0.04**	**0.0111**	**0.0076**

All significant loci results (*p<0.05) for combined UCSF + ADGC cohort (n = 11,690) when males and females are analyzed separately. Alleles present in one of the top three most significant 5-allele haplotypes in the full cohort (02:01~13:02~07:01~02:01~02:02, 03:01~07:02~15:01~01:02~06:02 and 24:02~38:01~13:01~01:03~06:03) are shown in this table in bold. Non-significant results are shown in grey and significant findings are shown in bold. In addition to odds ratio (OR) with 95% confidence interval (CI), a breakdown of allele frequency in Alzheimer's disease cases versus healthy older adult controls is also provided.

Iterative subanalyses corroborate role of HLA-A*03:01~B*07:02~DRB1*15:01~DQA1*01:02~DQB1*06:02 in AD

To attempt to alleviate concern over possible Type I error in this analysis, we randomly split the combined ADGC+UCSF cohort ten times (maintaining the same proportion of AD:controls) and reran the five-allele haplotype analysis in the 20 resulting (smaller) cohorts. Two of the top-associated five-allele haplotypes showed *p*-values < 0.05 in over half of the randomly split analyses (Table 11), which was more than any of the other "top" haplotypes from the original analysis. This included the one we focused on in this study (*A*03:01~B*07:02~DRB1*15:01~DQA1*01:02~DQB1*06:02*), which showed significance in 11 iterations of the randomly split analysis, with *p*-values from 0.026–0.0001 and ORs of 1.21–1.40, further corroborating the contributions of this haplotype and its subcomponents to AD risk.

Table 11. Twenty iterative analyses of randomly-split ADGC and UCSF combined cohort to corroborate top 5-allele haplotype associations with Alzheimer's disease risk.

Haplotype A~B~DRB1~DQA1~DQB1	OR (95% CI)	Original *p*-Value	*n* rep	%
01:01~08:01~07:01~02:01~03:03	0.43 (0.16 - 1.02)	0.037	5	25
02:01~13:02~07:01~02:01~02:02	**0.66 (0.50 - 0.89)**	**0.004**	**11**	**55**
02:01~15:01~07:01~02:01~02:02	0.39 (0.14 - 0.99)	0.030	6	30
02:01~44:02~13:01~01:03~06:03	1.44 (1.03 - 2.03)	0.027	6	30
02:01~57:01~07:01~02:01~03:03	1.31 (1.01 - 1.69)	0.038	6	30
03:01~07:02~12:01~05:05~03:01	0.30 (0.09 - 0.84)	0.011	1	5
03:01~07:02~15:01~01:02~06:02	**1.22 (1.08 - 1.38)**	**0.001**	**11**	**55**
11:01~35:01~07:01~02:01~02:02	0.31 (0.07 - 1.01)	0.031	0	0
24:02~38:01~13:01~01:03~06:03	**0.14 (0.02 - 0.63)**	**0.003**	**4**	**20**
24:02~44:05~01:01~01:01~05:01	4.56 (0.94 - 43.38)	0.033	0	0
29:02~58:01~08:04~04:01~04:02	4.56 (0.94 - 43.38)	0.033	0	0
68:01~44:02~01:01~01:01~05:01	1.96 (0.99 - 4.04)	0.038	7	35

We randomly split our full ADGC + UCSF cohort (n = 11,381) in half 10 times, balancing cases and controls, to determine how often the haplotypes we found to be significant ($p < 0.05$) in our original analysis replicated in these 20 smaller cohort analyses. Odds ratio (OR), 95% confidence interval (CI), and p-value from original analysis (all values from Table 3) are listed for each haplotype, and results are listed for any of the replication cohorts in which a given haplotype was significant. Each haplotype also lists the number of smaller cohorts (n rep) and the percentage (number of replication cohorts / 20) in which a significant finding was found. The three most significant findings in our original 5-allele haplotype analysis in the combined ADGC + UCSF cohort are highlighted here in blue.

Discussion

In a total of over 11,000 individuals, we found evidence suggesting that the five-allele HLA haplotype *A*03:01~B*07:02~DRB1*15:01~DQA1*01:02~DQB1*06:02* is a risk factor for AD and that this effect may be driven by men who do not carry the major AD risk factor, *APOE-ε4*. Locus-level analysis further confirmed AD associations of the individual alleles *B*07:02*, *DRB1*15:01*, *DQA1*01:02*, and *DQB1*06:02*. In separate class I and class II haplotype analyses, the class I *A*03:01~B*07:02* haplotype and the class II *DRB1*15:01~DQA1*01:02~DQB1*06:02* (*DR15*) haplotype were both significantly associated with risk for AD. We assessed the clinical relevance of each of these haplotypes separately in a smaller cohort representing the spectrum of cognitively normal controls and individuals with MCI and AD. Carrying the MHC class I haplotype *A*03:01~B*07:02* was associated with higher CSF amyloid levels, suggesting lower levels of amyloid in the brains of haplotype carriers across the AD spectrum. The class II haplotype *DR15* was associated with greater rate of decline on two different measures of cognitive function relevant to AD in a dose-dependent manner. In a subset of the same cohort, carrying the *DR15* risk haplotype was also associated with higher baseline levels of CC4, a biomarker of AD-related inflammation [46]. Taking these findings together, this study provides evidence for the contribution of the *A*03:01~B*07:02 ~DRB1*15:01~DQA1*01:02~DQB1*06:02* haplotype and its components, *A*03:01 ~B*07:02* and *DR15*, to risk of AD.

Over 30 years of research into HLA alleles and risk of AD has yielded mixed conclusions due in part to limitations in mapping alleles within this complicated genomic region. Early studies mapped risk of AD to the HLA region of chromosome 6 [47], and

the studies that followed differed significantly in their methodological approach, the identities and resolution of the alleles studied, the ethnicity of the study cohorts, and the inferences drawn from the data. MHC class I molecule *HLA-A*02* has been shown to either be associated with increased risk of AD or to have no effect in nearly 15 different studies [48–61]. Given that only *B*07:02*, *DRB1*15:01*, *DQA1*01:02*, and *DQB1*06:02* showed significant locus-level associations with AD, our findings are consistent with an ambiguous role of *HLA-A*02* in AD. In terms of class II alleles, one study by Mansouri and colleagues demonstrated a link between *DRB1*15:01~DQB1*06:02* and AD in a small cohort of Tunisians [62], consistent with our findings. Previous GWASs have found that AD risk is associated with a SNP in *DRB5* [63]. As there is strong LD between *DRB5*01* and *DRB1*15:01~DQB1*06:02* [64], it is possible that the AD association we have detected with *DR15* is due in part or wholly to *DRB5*. Finally, our finding that HLA associations with AD are stronger in *APOE-ε4* -negative individuals is consistent with prior work for different HLA alleles [65,66].

The HLA region has been studied to a varying extent for its contributions to neurological disease, and many of the risk alleles implicated in the present studies have also been linked to other disorders. Most notably, the class II *DR15* haplotype is the most consistently replicated genetic finding in MS [67–69]. *DR15* also correlates with worse clinical progression in women with relapsing-onset MS (e.g., younger age at onset and more subcortical atrophy) [44]. Class I allele *B*07* has also been associated with MS risk, particularly in those also carrying *DRB1*15* [70]. In one Parkinson's disease (PD) study, four alleles identified in a risk haplotype overlapped with our top five-allele haplotype association [71]. Similar to AD, other studies have also implicated

the *HLA-DRB5* region in PD risk [72]. In one small autism study, the class I allele *B*07* and class II allele *DQB1*06:02* were both associated with disease risk [73]. Finally, class II allele *DQB1*06:02* has been associated with marked increased risk for [74,75], and worsened severity of [76,77], narcolepsy. These findings are consistent with our study, in which we identified a dose-dependent association between *DR15* and greater cognitive decline in individuals representing the AD spectrum.

Participants who carried at least one copy of the class I haplotype *A*03:01~B*07:02* on average had higher baseline CSF amyloid levels, suggesting lower amyloid burden in the brains of these participants. Similar findings have been observed in *APOE-ε4*-negative individuals when compared to *APOE-ε4*-positive individuals across the phenotypic spectrum of cognitively normal to early MCI, suggesting higher brain amyloid β in carriers [78]. This finding raises the possibility that there could be a tau-mediated effect on AD clinical symptoms, as the AD group did not differ in clinical measures by haplotype carrier status. On the other hand, *DR15* haplotype carriers demonstrated subtle differences in baseline inflammatory biomarker levels, as well as a worse cognitive trajectory over time, suggesting a disease-modifying effect that could be mediated by changes in immune function.

Our study benefited from several strengths. The primary discovery cohort was a well-characterized sample of patients who received extensive clinical evaluations at the UCSF Memory and Aging Center. The replication dataset from the ADGC of over 11,000 AD and cognitively normal control individuals is the largest dataset to date used to explore immunogenetic contributions to AD risk. Lastly, longitudinal data from ADNI allowed us to probe the potential clinical relevance of haplotype findings across the AD

spectrum. Our study also has caveats that are important to consider. Our imputation program predicted accuracy for individuals with European ancestry that was likely higher than 94.8%–99.2% based on our more stringent call threshold in comparison to other studies [38]. Imputed HLA alleles have been shown to be reliable classification tools in studies with similar methodologies [38,79,80]; however, future studies would benefit from direct sequencing of HLA alleles to avoid potential imputation inaccuracies. Because of limitations in the imputation package selected for HLA allele calling, we were only able to impute genotypes for a subset of MHC class I and II genes. For example, the imputed genes available did not include *DRB5*, which was indicated in previous studies to be associated with AD risk or pathological processes [9,24,81]. *DRB5* is on the *DR15* haplotype, so it is likely that the association we identified reflects these previous results. However, we can neither directly confirm nor refute this possibility in the present study. The *DR15* risk haplotype is most common in Europeans, and to minimize genetic heterogeneity in population substructure, we limited the present analysis to white individuals of non-Hispanic descent. Additional studies are required to assess the identified HLA risk haplotypes and component alleles for their contribution to AD in more diverse populations where patterns of LD differ and may uncouple alleles that were tightly linked in our study population, though the initial study identifying class II associations with AD in Tunisians suggests this may be a generalized risk phenomenon. Although $p < 0.05$ may be considered lenient based on the number of total alleles tested, it is also true that all of these alleles represent only five genes within one genomic region that has been previously linked to AD risk. Despite reduced statistical power due to low frequency of HLA haplotypes imparted by the extraordinary

diversity of this region, we feel that this study is an important first step in elucidating the underlying contribution of the HLA to AD risk given the medical implications of ultimately identifying immune-related therapies as a means of modifying a complex, common disease. We have greater confidence in our findings due to corroborating clinical validity as identified in the ADNI cohort. Iterative subanalyses of the combined study cohort further support a role of our top five-allele haplotypes in AD risk. In addition, two of the main alleles of interest we identified, *DRB1*15:01* and *DQB1*06:02*, have been linked to AD risk in two prior studies, further supporting our results. We also identified several other risk haplotypes in our analyses beyond the ones we focused on in this study; the clinical relevance of these additional haplotypes and alleles requires further investigation. Future work is also required to test whether these findings extend to early-onset and atypical clinical syndromes with underlying AD pathology.

In summary, we present evidence for a role of the HLA class I *A*03:01~B*07:02* haplotype and the HLA class II *DR15* haplotype in AD risk. Our study also suggests that these risk haplotypes may be associated with CSF AD biomarker levels (class I) and greater decline in cognition over time, as well as higher levels of inflammation across aging (class II). The results of our study indicate that the broad *A*03:01~B*07:02 ~DRB1*15:01~DQA1*01:02~DQB1*06:02* haplotype may contribute genetic risk to AD beyond that contributed by the established risk factor *APOE-ε4*, particularly in men. As components of this haplotype are well-established risk factors in MS, PD, autism, and narcolepsy, we propose that they may contribute to underlying biological risk mechanisms in multiple neurological diseases. Future work is required to establish the precise molecular processes underlying this risk association, as well as to expand this

finding to broader, diverse populations of AD and potentially even other neurodegenerative conditions.

Acknowledgments

We thank all participants and their families for their contributions to research. We thank Danny Park for his consultation on statistical approaches. Samples from the National Cell Repository for Alzheimer's Disease (NCRAD) were used in this study. We thank contributors who collected samples used in this study, as well as patients and their families, whose help and participation made this work possible. We acknowledge the Alzheimer's Disease Genetic Consortium for the collection and genotyping of samples used in this study. Full ADGC and ADNI acknowledgments are provided in S1 Acknowledgments.

References

1. Prince M, Wimo A, Guerchet M, Ali G-C, Wu Y-T, Prina M, et al. World Alzheimer Report 2015. The global impact of dementia: An analysis of prevalence, incidence, cost and trends. 2015. https://www.alz.co.uk/research/world-report-2015

2. Zetzsche T, Rujescu D, Hardy J, Hampel H. Advances and perspectives from genetic research: development of biological markers in Alzheimer's disease. Expert Rev Mol Diagn. 2010;10: 667–690.

3. Guerreiro R, Wojtas A, Bras J, Carrasquillo M, Rogaeva E, Majounie E, et al. TREM2 Variants in Alzheimer's Disease. N Engl J Med. 2013;368(2): 117–127.

4. Jonsson T, Stefansson H, Steinberg S, Jonsdottir I, Jonsson P, Snaedal J, et al. Variant of TREM2 Associated with the Risk of Alzheimer's Disease. N Engl J Med. 2013;368(2): 107–116.

5. Gonzalez Murcia JD, Schmutz C, Munger C, Perkes A, Gustin A, Peterson M, et al. Assessment Of Trem2 Rs75932628 Association With Alzheimer's Disease In A Population-Based Sample: The Cache County Study. Neurobiol Aging. 2013;34: 2889.e11-2889.e13. doi:10.1016/j.neurobiolaging.2013.06.004

6. Sirkis DW, Bonham LW, Aparicio RE, Geier EG, Ramos EM, Wang Q, et al. Rare TREM2 variants associated with Alzheimer's disease display reduced cell surface expression. Acta Neuropathol Commun. 2016;4(1):98.

7. Benitez BA, Cooper B, Pastor P, Jin SC, Lorenzo E, Cervantes S, et al. TREM2 is associated with risk of Alzheimer disease in Spanish population. Neurobiol Aging. 2013;34: 1711.e15-1711.e17. doi:10.1016/j.neurobiolaging.2012.12.018

8. Coppola G, Chinnathambi S, Lee JJ, Dombroski BA, Baker MC, Soto-Ortolaza AI, et al. Evidence for a role of the rare p.A152T variant in MAPT in increasing the risk for FTD-spectrum and Alzheimer's diseases. Hum Mol Genet. 2012;21: 3500–3512.

9. Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet. 2013;45: 1452–8. doi:10.1038/ng.2802

10. Naj AC, Jun G, Beecham GW, Wang L, Vardarajan BN, Buros J, et al. Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. Nat Genet. 2011;43: 436–441.

11. Karch CM, Goate AM. Alzheimer's disease risk genes and mechanisms of disease pathogenesis. Biol Psychiatry. 2015;77: 43–51. doi:10.1016/j.biopsych.2014.05.006

12. Hensley K. Neuroinflammation in Alzheimer's disease: mechanisms, pathologic consequences, and potential for therapeutic manipulation. J Alzheimers Dis. 2010;21: 1–14. doi:10.3233/JAD-2010-1414

13. Wes PD, Holtman IR, Boddeke EWGM, Möller T, Eggen BJL. Next generation transcriptomics and genomics elucidate biological complexity of microglia in health and disease. Glia. 2016;64: 197–213. doi:10.1002/glia.22866

14. Gan L, Ye S, Chu A, Anton K, Yi S, Vincent VA, et al. Identification of Cathepsin B as a Mediator of Neuronal Death Induced by Aβ-activated Microglial Cells Using a Functional Genomics Approach. J Biol Chem. 2004;279: 5565–5572.

15. Hong S, Beja-Glasser VF, Nfonoyim BM, Frouin A, Li S, Ramakrishnan S, et al. Complement and microglia mediate early synapse loss in Alzheimer mouse models. Science. 2016;352: 712–6. doi:10.1126/science.aad8373

16. Eikelenboom P, Hoozemans JJ, Veerhuis R, van Exel E, Rozemuller AJ, van Gool WA. Whether, when and how chronic inflammation increases the risk of developing late-onset Alzheimer's disease. Alzheimers Res Ther. 2012;4: 15. doi:10.1186/alzrt118

17. Schmidt R, Schmidt H, Curb JD, Masaki K, White LR, Launer LJ. Early inflammation and dementia: A 25-year follow-up of the Honolulu-Asia Aging Study. Ann Neurol. 2002;52: 168–174. doi:10.1002/ana.10265

18. Mrak RE, Griffin WST. Potential inflammatory biomarkers in Alzheimer's disease. J Alzheimers Dis. 2005;8: 369–75.

19. Wood JA, Wood PL, Ryan R, Graff-Radford NR, Pilapil C, Robitaille Y, et al. Cytokine indices in Alzheimer's temporal cortex: no changes in mature IL-1 beta or IL-1RA but increases in the associated acute phase proteins IL-6, alpha 2-macroglobulin and C-reactive protein. Brain Res. 1993;629: 245–52.

20. Shi Q, Colodner KJ, Matousek SB, Merry K, Hong X, Kenison JE, et al. Complement C3-Deficient Mice Fail to Display Age-Related Hippocampal Decline. J Neurosci. 2015;35: 13029–13042. doi:10.1523/JNEUROSCI.1698-15.2015

21. Stephan AH, Madison D V, Mateos JM, Fraser DA, Lovelett EA, Coutellier L, et al. A Dramatic Increase of C1q Protein in the CNS during Normal Aging. J Neurosci. 2013;33: 13460–13474. doi:10.1523/JNEUROSCI.1333-13.2013

22. Dorman JS, Bunker CH. HLA-DQ locus of the human leukocyte antigen complex and type 1 diabetes mellitus: a HuGE review. Epidemiol Rev. 2000;22: 218–27.

23. Allen M, Kachadoorian M, Carrasquillo MM, Karhade A, Manly L, Burgess JD, et al. Late-onset Alzheimer disease risk variants mark brain regulatory loci. Neurol Genet. 2015;1: e15. doi:10.1212/NXG.0000000000000012

24. Yokoyama JS, Wang Y, Schork AJ, Thompson WK, Karch CM, Cruchaga C, et al. Association Between Genetic Traits for Immune-Mediated Diseases and Alzheimer Disease. JAMA Neurol. 2016;94158: 1–7. doi:10.1001/jamaneurol.2016.0150

25. Dubois B, Feldman H, Jacova C. Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria. Lancet Neurol. 2014;13: 614–29. doi:10.1016/S1474-4422(14)70090-0

26. Boehme KL, Young B, Mukherjee S, Crane PK, Kauwe J, Young B. ADGC 1000 Genomes combined data workflow. 2014; 1–12. http://kauwelab.byu.edu/Portals/22/adgc_combined_1000G_12032014.pdf

27. Bonham LW, Desikan RS, Yokoyama JS. The relationship between complement factor C3, APOE ε4, amyloid and tau in Alzheimer's disease. Acta Neuropathol Commun. 2016;4: 1–7. doi:10.1186/s40478-016-0339-y

28. Desikan RS, Thompson WK, Holland D, Hess CP, Brewer JB, Zetterberg H, et al. The role of clusterin in amyloid-β-associated neurodegeneration. JAMA Neurol. 2014;71: 180–7. doi:10.1001/jamaneurol.2013.4560

29. Desikan RS, Thompson WK, Holland D, Hess CP, Brewer JB, Zetterberg H, et al. Heart fatty acid binding protein and Aβ-associated Alzheimer's neurodegeneration. Mol Neurodegener. 2013;8: 39. doi:10.1186/1750-1326-8-39

30. Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. Neurology. 1993;43: 2412–2414. doi:10.1212/WNL.43.11.2412-a

31. Saykin AJ, Shen L, Foroud TM, Potkin SG, Swaminathan S, Kim S, et al. Alzheimer's Disease Neuroimaging Initiative biomarkers as quantitative phenotypes: Genetics core aims, progress, and plans. Alzheimers Dement. 2010;6(3): 265–273. doi:10.1016/j.jalz.2010.03.013

32. Shaw LM, Vanderstichele H, Knapik-Czajka M, Clark CM, Aisen PS, Petersen RC, et al. Cerebrospinal fluid biomarker signature in alzheimer's disease neuroimaging initiative subjects. Ann Neurol. 2009;65: 403–413. doi:10.1002/ana.21610

33. Rey A. L'examen clinique en psychologie. Lexamen clinique en psychologie. 1958. doi:CRJ Alz lib Copy2 #618; CRJ Alz lib Copy2-Converted #618; CRJ Alz Library 2 shared lib-Converted #1017; MMM Master_Reference_List-Converted #21; JLW MCI-Converted #22; JLW PCA-Converted #58; JLW VBMprediction-Converted #87

34. Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer's disease. Am J Psychiatry. 1984;141(11): 1356–1364. doi:10.1176/ajp.141.11.1356

35. Mohs RC CL. Alzheimer's Disease Assessment Scale (ADAS). Psychopharmacol Bull. 1988;24: 627–8.

36. Hughes CP, Berg L, Danziger WL. A new clinical scale for the staging of dementia. Br J Psychiatry. 1982;140: 566–572. doi:10.1192/bjp.140.6.566

37. Khor S-S, Yang W, Kawashima M, Kamitsuji S, Zheng X, Nishida N, et al. High-accuracy imputation for HLA class I and II genes based on high-resolution SNP data of population-specific references. Pharmacogenomics J. 2015;15: 530–7. doi:10.1038/tpj.2015.4

38. Zheng X, Shen J, Cox C, Wakefield JC, Ehm MG, Nelson MR, et al. HIBAG—HLA genotype imputation with attribute bagging. Pharmacogenomics J. 2014;14: 192–200. doi:10.1038/tpj.2013.18

39. Pappas DJ, Marin W, Hollenbach JA, Mack SJ. Bridging ImmunoGenomic Data Analysis Workflow Gaps (BIGDAWG): An integrated case-control analysis pipeline. Hum Immunol. 2016;77: 283–7. doi:10.1016/j.humimm.2015.12.006

40. Van Deerlin VM, Sleiman PM, Martinez-Lage M, Chen-Plotkin A, Wang L-S, Graff-Radford NR, et al. Common variants at 7p21 are associated with frontotemporal lobar degeneration with TDP-43 inclusions. Nat Genet. 2010;42: 234–9. doi:10.1038/ng.536

41. Hollenbach JA, Mack SJ, Thomson G, Gourraud PA. Analytical methods for disease association studies with immunogenetic data. Methods Mol Biol. 2012;882: 245–266. doi:10.1007/978-1-61779-842-9_14

42. Maiers, M., Gragert, L., Klitz, W. High resolution HLA alleles and haplotypes in the US population. Human Immunology (2007) 68, 779-788.

43. Tapiola T, Pirttilä T, Mehta PD, Alafuzof I, Lehtovirta M, Soininen H. Relationship between apoE genotype and CSF β-amyloid (1-42) and tau in patients with probable and definite Alzheimer's disease. Neurobiol Aging. 2000;21: 735–740. doi:10.1016/S0197-4580(00)00164-0

44. Isobe N, Keshavan A, Gourraud P-A, Zhu AH, Datta E, Schlaeger R, et al. Association of HLA Genetic Risk Burden With Disease Phenotypes in Multiple Sclerosis. JAMA Neurol. 2016;73: 795–802. doi:10.1001/jamaneurol.2016.0980

45. Farrer LA, Cupples LA, Haines JL, Hyman B, Kukull WA, Mayeux R, et al. Effects of Age, Sex, and Ethnicity on the Association Between Apolipoprotein E Genotype and Alzheimer Disease: A Meta-analysis. JAMA. 1997;278(16): 1349–1356. doi:10.1001/jama.1997.03550160069041

46. Kessler H, Pajonk F-G, Meisser P, Schneider-Axmann T, Hoffmann K-H, Supprian T, et al. Cerebrospinal fluid diagnostic markers correlate with lower plasma copper and ceruloplasmin in patients with Alzheimer's disease. J Neural Transm. 2006;113: 1763–9. doi:10.1007/s00702-006-0485-7

47. Weitkamp LR, Nee L, Keats B, Polinsky RJ, Guttormsen S. Alzheimer disease: evidence for susceptibility loci on chromosomes 6 and 14. Am J Hum Genet. 1983;35: 443–53.

48. Payami H, Schellenberg GD, Zareparsi S, Kaye J, Sexton GJ, Head MA, et al. Evidence for association of HLA-A2 allele with onset age of Alzheimer's disease. Neurology. 1997;49: 512–8.

49. Payami H, Kaye J, Becker W, Norman D, Wetzsteon P. HLA-A2, or a closely linked gene, confers susceptibility to early-onset sporadic Alzheimer's disease in men. Neurology. 1991;41: 1544–8.

50. Combarros O, Escribano J, Sánchez-Velasco P, Leyva-Cobián F, Oterino A, Leno C, et al. Association of the HLA-A2 allele with an earlier age of onset of Alzheimer's disease. Acta Neurol Scand. 1998;98: 140–1.

51. Guerini FR, Calabrese E, Agliardi C, Zanzottera M, Franceschi M, Grimaldi LM, et al. Association study of the HLA-A2 allele in Italian Alzheimer disease patients. Neurobiol Aging. 2009;30: 2082–2083.

52. Listì F, Candore G, Balistreri CR, Grimaldi MP, Orlando V, Vasto S, et al. Association between the HLA-A2 allele and Alzheimer disease. Rejuvenation Res. 2006;9: 99–101. doi:10.1089/rej.2006.9.99

53. Ma SL, Tang NLS, Tam CWC, Lui VWC, Suen EWC, Chiu HFK, et al. Association between HLA-A alleles and Alzheimer's disease in a southern Chinese community. Dement Geriatr Cogn Disord. 2008;26: 391–7. doi:10.1159/000164275

54. Araria-Goumidi L, Lambert JC, Cottel D, Amouyel P, Chartier-Harlin MC. No association of the HLA-A2 allele with Alzheimer's disease. Neurosci Lett. 2002;335: 75–8.

55. Zareparsi S, James DM, Kaye JA, Bird TD, Schellenberg GD, Payami H. HLA-A2 homozygosity but not heterozygosity is associated with Alzheimer disease. Neurology. 2002;58: 973–5.

56. Harris JM, Cumming AM, Craddock N, St Clair D, Lendon CL. Human leucocyte antigen-A2 increases risk of Alzheimer's disease but does not affect age of onset in a Scottish population. Neurosci Lett. 2000;294: 37–40.

57. Small GW, Scott WK, Komo S, Yamaoka LH, Farrer LA, Auerbach SH, et al. No association between the HLA-A2 allele and Alzheimer disease. Neurogenetics. 1999;2: 177–82.

58. Small GW, Ebeling SC, Matsuyama SS, Heyman A, Reisner EG, Renvoize EB, et al. Variable association of HLA-A2 in men with early-onset Alzheimer disease. Neurobiol Aging. 1991;12: 375–7.

59. Ballerini C, Nacmias B, Rombolà G, Marcon G, Massacesi L, Sorbi S. HLA A2 allele is associated with age at onset of Alzheimer's disease. Ann Neurol. 1999;45: 397–400.

60. Guerini FR, Tinelli C, Calabrese E, Agliardi C, Zanzottera M, De Silvestri A, et al. HLA-A*01 is associated with late onset of Alzheimer's disease in Italian patients. Int J Immunopathol Pharmacol. 2009;22: 991–9.

61. Middleton D, Mawhinney H, Curran MD, Edwardson JA, Perry R, McKeith I, et al. Frequency of HLA-A and B alleles in early and late-onset Alzheimer's disease. Neurosci Lett. 1999;262: 140–2.

62. Mansouri L, Messalmani M, Klai S, Bedoui I, Derbali H, Gritli N, et al. Association of HLA-DR/DQ polymorphism with Alzheimer's disease. Am J Med Sci. 2015;349: 334–7. doi:10.1097/MAJ.0000000000000416

63. Yokoyama JS, Wang Y, Schork AJ, Thompson WK, Karch CM, Cruchaga C, et al. Association between genetic traits for immune-mediated diseases and Alzheimer disease. JAMA Neurol. 2016;73. doi:10.1001/jamaneurol.2016.0150

64. Miretti MM, Walsh EC, Ke X, Delgado M, Griffiths M, Hunt S, et al. A High-Resolution Linkage-Disequilibrium Map of the Human Major Histocompatibility Complex and First Generation of Tag Single-Nucleotide Polymorphisms. Am J Hum Genet. 2005;76: 634–646. doi:10.1086/429393

65. Lehmann DJ, Wiebusch H, Marshall SE, Johnston C, Warden DR, Morgan K, et al. HLA class I, II & III genes in confirmed late-onset Alzheimer's disease. Neurobiol Aging. 2001;22: 71–77.

66. Neill D, Curran MD, Middleton D, Mawhinney H, Edwardson JA, McKeith I, et al. Risk for Alzheimer's disease in older late-onset cases is associated with HLA-DRB1*03. Neurosci Lett. 1999;275: 137–40.

67. Schmidt H, Williamson D, Ashley-Koch A. HLA-DR15 haplotype and multiple sclerosis: a HuGE review. Am J Epidemiol. 2007;165: 1097–109. doi:10.1093/aje/kwk118

68. International Multiple Sclerosis Genetics Consortium, Wellcome Trust Case Control Consortium 2, Sawcer S, Hellenthal G, Pirinen M, Spencer CCA, et al. Genetic risk and

a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature. 2011;476: 214–9. doi:10.1038/nature10251

69. Patsopoulos NA, Barcellos LF, Hintzen RQ, Schaefer C, van Duijn CM, Noble JA, et al. Fine-mapping the genetic association of the major histocompatibility complex in multiple sclerosis: HLA and non-HLA effects. PLoS Genet. 2013;9: e1003926. doi:10.1371/journal.pgen.1003926

70. Lysandropoulos AP, Racapé J, Holovska V, Toungouz M. Human leucocyte antigen (HLA) class I and II typing in Belgian multiple sclerosis patients. Acta Neurol Belg. 2016; doi:10.1007/s13760-016-0716-0

71. Wissemann WT, Hill-Burns EM, Zabetian CP, Factor SA, Patsopoulos N, Hoglund B, et al. Association of Parkinson disease with structural and regulatory variants in the HLA region. Am J Hum Genet. 2013;93: 984–993. doi:10.1016/j.ajhg.2013.10.009

72. Nalls MA, Plagnol V, Hernandez DG, Sharma M, Sheerin UM, Saad M, et al. Imputation of sequence variants for identification of genetic risks for Parkinson's disease: A meta-analysis of genome-wide association studies. Lancet. 2011;377: 641–649. doi:10.1016/S0140-6736(10)62345-8

73. Al-Hakbany M, Awadallah S, AL-Ayadhi L. The Relationship of HLA Class I and II Alleles and Haplotypes with Autism: A Case Control Study. Autism Res Treat. 2014;2014: 1–6. doi:10.1155/2014/242048

74. Tafti M, Hor H, Dauvilliers Y, Lammers GJ, Overeem S, Mayer G, et al. DQB1 Locus Alone Explains Most of the Risk and Protection in Narcolepsy with Cataplexy in Europe. Sleep. 2014;37: 19–25. doi:10.5665/sleep.3300

75. Chabas D, Taheri S, Renier C, Mignot E. The genetics of narcolepsy. Annu Rev Genomics Hum Genet. 2003;4: 459–83. doi:10.1146/annurev.genom.4.070802.110432

76. Mignot E, Hayduk R, Black J, Grumet FC, Guilleminault C. HLA DQB1*0602 is associated with cataplexy in 509 narcoleptic patients. Sleep. 1997;20: 1012–20.

77. Watson NF, Ton TGN, Koepsell TD, Gersuk VH, Longstreth WT, Jr. Does narcolepsy symptom severity vary according to HLA-DQB1*0602 allele status? Sleep. 2010;33: 29–35.

78. Risacher SL, Kim S, Nho K, Foroud T, Shen L, Petersen RC, et al. APOE effect on Alzheimer's disease biomarkers in older adults with significant memory concern. Alzheimer's Dement. 2015;11: 1417–1429. doi:10.1016/j.jalz.2015.03.003

79. Hsieh A-R, Chang S-W, Chen P-L, Chu C-C, Hsiao C-L, Yang W-S, et al. Predicting HLA genotypes using unphased and flanking single-nucleotide polymorphisms in Han Chinese population. BMC Genomics. 2014;15: 81. doi:10.1186/1471-2164-15-81

80. Dilthey A, Leslie S, Moutsianas L, Shen J, Cox C, Nelson MR, et al. Multi-Population Classical HLA Type Imputation. PLoS Comput Biol. 2013;9. doi:10.1371/journal.pcbi.1002877

81. Yu L, Chibnik LB, Srivastava GP, Pochet N, Yang J, Xu J, et al. Association of Brain DNA methylation in SORL1, ABCA7, HLA-DRB5, SLC24A4, and BIN1 with pathological

diagnosis of Alzheimer disease. JAMA Neurol. 2015;72: 15–24. doi:10.1001/jamaneurol.2014.3049

Chapter II: Appendix

S1 Acknowledgments.

Detailed acknowledgments for ADNI and ADGC and other samples used in this study.

ADNI Cohort Recognition

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and
DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

ADGC

The National Institutes of Health, National Institute on Aging (NIH-NIA) supported this work through the following grants: ADGC, U01 AG032984, RC2 AG036528; Samples from the National Cell Repository for Alzheimer's Disease (NCRAD), which receives government support under a cooperative agreement grant (U24 AG21886) awarded by the National Institute on Aging (NIA), were used in this study. We thank contributors who collected samples used in this study, as well as patients and their families, whose help and participation made this work possible; Data for this study were prepared, archived, and distributed by the National Institute on Aging Alzheimer's Disease Data Storage Site (NIAGADS) at the University of Pennsylvania (U24-AG041689-01); NACC, U01 AG016976; NIA LOAD, U24 AG026395, R01AG041797; Banner Sun Health Research Institute P30 AG019610; Boston University, P30 AG013846, U01 AG10483, R01 CA129769, R01 MH080295, R01 AG017173, R01 AG025259, R01AG33193; Columbia University, P50 AG008702, R37 AG015473; Duke University, P30 AG028377, AG05128; Emory University, AG025688; Group Health Research Institute, UO1 AG006781, UO1 HG004610, UO1 HG006375; Indiana University, P30 AG10133; Johns Hopkins University, P50 AG005146, R01 AG020688; Massachusetts General Hospital, P50 AG005134; Mayo Clinic, P50 AG016574; Mount Sinai School of Medicine, P50 AG005138, P01 AG002219; New York University, P30 AG08051, UL1 RR029893, 5R01AG012101, 5R01AG022374, 5R01AG013616, 1RC2AG036502, 1R01AG035137; Northwestern University, P30 AG013854; Oregon Health & Science University, P30 AG008017, R01 AG026916; Rush University, P30 AG010161, R01 AG019085, R01 AG15819, R01 AG17917, R01 AG30146; TGen, R01 NS059873; University of Alabama at Birmingham, P50 AG016582; University of Arizona, R01 AG031581; University of California, Davis, P30 AG010129; University of California, Irvine, P50 AG016573; University of California, Los Angeles, P50 AG016570; University of California, San Diego, P50 AG005131; University of California, San Francisco, P50 AG023501, P01 AG019724; University of Kentucky, P30 AG028383, AG05144; University of Michigan, P50 AG008671; University of Pennsylvania, P30 AG010124; University of Pittsburgh, P50 AG005133, AG030653, AG041718, AG07562, AG02365; University of Southern California, P50 AG005142; University of Texas Southwestern, P30 AG012300; University of Miami, R01 AG027944, AG010491, AG027944, AG021547, AG019757; University of Washington, P50 AG005136; University of Wisconsin, P50 AG033514; Vanderbilt University, R01 AG019085; and Washington University, P50 AG005681, P01 AG03991. The Kathleen Price Bryan Brain Bank at Duke University Medical Center is funded by NINDS grant # NS39764, NIMH MH60451 and by Glaxo Smith Kline. Genotyping of the TGEN2 cohort was supported by

Kronos Science. The TGen series was also funded by NIA grant AG041232 to AJM and MJH, The Banner Alzheimer's Foundation, The Johnnie B. Byrd Sr. Alzheimer's Institute, the Medical Research Council, and the state of Arizona and also includes samples from the following sites: Newcastle Brain Tissue Resource (funding via the Medical Research Council, local NHS trusts and Newcastle University), MRC London Brain Bank for Neurodegenerative Diseases (funding via the Medical Research Council),South West Dementia Brain Bank (funding via numerous sources including the Higher Education Funding Council for England (HEFCE), Alzheimer's Research Trust (ART), BRACE as well as North Bristol NHS Trust Research and Innovation Department and DeNDRoN), The Netherlands Brain Bank (funding via numerous sources including Stichting MS Research, Brain Net Europe, Hersenstichting Nederland Breinbrekend Werk, International Parkinson Fonds, Internationale Stiching Alzheimer Onderzoek), Institut de Neuropatologia, Servei Anatomia Patologica, Universitat de Barcelona. ADNI data collection and sharing was funded by the National Institutes of Health Grant U01 AG024904 and Department of Defense award number W81XWH-12-2-0012. ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California. We thank Drs. D. Stephen Snyder and Marilyn Miller from NIA who are *ex-officio* ADGC members. Support was also from the Alzheimer's Association (LAF, IIRG-08-89720; MP-V, IIRG-05-14147) and the US Department of Veterans Affairs Administration, Office of Research and Development, Biomedical Laboratory Research Program. P.S.G.-H. is supported by Wellcome Trust, Howard Hughes Medical Institute, and the Canadian Institute of Health Research.

ROSMAP (subset of ADGC):
The ROSMAP dataset is supported by the following grants: P30AG10161, R01AG15819, R01AG17917.

NCRAD: Samples from the National Cell Repository for Alzheimer's Disease (NCRAD), which receives government support under a cooperative agreement grant (U24 AG21886) awarded by the National Institute on Aging (NIA), were used in this study.

S1 List. List of immune and inflammation-related CSF biomarkers included in analysis

Analyte
Beta 2 Microglobulin
C-Reactive Protein
CD 40 antigen
Chemokine CC-4
Complement C3
Cortisol
Fas Ligand
Immunoglobulin A
Intercellular Adhesion Molecule 1
Interferon gamma Induced Protein 10
Interleukin-16
Interleukin-25
Interleukin-3
Interleukin-6 receptor
Interleukin-8
Macrophage Colony-Stimulating Factor 1
Macrophage Inflammatory Protein-1 beta
Macrophage Migration Inhibitory Factor
Monocyte Chemotactic Protein 1
Monocyte Chemotactic Protein 2
Monokine Induced by Gamma Interferon
Neutrophil Gelatinase-Associated Lipocalin
Osteopontin
Resistin
Stem Cell Factor
T Lymphocyte-Secreted Protein I-309
T-Cell-Specific Protein RANTES
Thrombomodulin

S1 Methods. ADNI Diagnosis and CSF Biomarker Information

ADNI Diagnostic Guidelines

Controls were required to have normal memory function on the Logical Memory II subscale of the Weschler Memory Scale– Revised (Binder, Storandt, & Birge, 1999), an MMSE score greater than 24, CDR total score equal to 0, and clinical determination that the individual was not significantly impaired in cognitive function or activities of daily living. Individuals with MCI were required to have abnormal memory function on the Logical Memory II subscale of the Weschler Memory Scale – Revised, an MMSE greater than 24, CDR total score equal to 0.5, and clinical determination that the individual's general cognition and functional performance was impaired enough to make a diagnosis of AD. Finally, individuals with AD were required to have abnormal memory function on the Logical Memory II subscale of the Weschler Memory Scale – Revised, an MMSE between 20 and 26, CDR total score equal to 0.5 or 1.0, and judgment by a clinician that the individual met NINCDS/ADRDA criteria for probable AD (Mckhann et al., 2011).

ADNI CSF Biomarker Measurement

Baseline CSF CC4 levels were measured using the Human DiscoveryMAP panel developed by Rules Based Medicine (Myriad RBM; Austin, Texas). The Human DiscoveryMAP panel is commercially available and measures a collection of metabolic, lipid, inflammatory, and other AD-relevant indicators. A full list of the measured metabolites is available through Myriad RBM. The CSF measurements in the immunoassay panel were processed and normalized according to previously described methods (Craig-Schapiro et al., 2011; Siuciak, 2011). Briefly, Myriad RBM used a Luminex 100 instrument for the measurements and analyzed the resulting data using proprietary software. The ADNI staff checked analyte distributions for normality using Box-Cox analyses and, if needed, log10 transformed the data to achieve an approximately normal distribution. Out of 83 biomarkers with sufficient data available for analysis, we selected 28 that were most directly relevant to immune function and inflammation (S1 List).

References

Binder, E. F., Storandt, M., & Birge, S. J. (1999). The Relation Between Psychometric Test Performance and Physical Performance in Older Adults. The Journals of Gerontology Series A: Biological Sciences and Medical Sciences, 54(8), M428–M432. https://doi.org/10.1093/gerona/54.8.M428

Craig-Schapiro, R., Kuhn, M., Xiong, C., Pickering, E. H., Liu, J., Misko, T. P., … Holtzman, D. M. (2011). Multiplexed immunoassay panel identifies novel CSF biomarkers for alzheimer's disease diagnosis and prognosis. PLoS ONE, 6(4). https://doi.org/10.1371/journal.pone.0018850

Mckhann, G., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R., Kawas, C., … Phelps, C. (2011). The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimer's and Dementia, 7(3), 263–269. JOUR.

Siuciak J. Biomarkers Consortium Data Primer. 2011

S1 Table. Summary information for datasets that make up ADGC merged dataset.

Study	Sex(M/F)	Cases/Controls/Missing	Sample Size
ACT 1	886/1,161	479/1,348/220	2,047
ADC 1	947/1,137	1,503/543/38	2,084
ADC 2	328/364	546/121/25	692
ADC 3	593/721	711/464/139	1,314
ADNI	317/207	215/140/169	524
GSK/GenADA	608/952	796/764/0	1,560
LOAD/NIA-LOAD	628/1,069	745/801/151	1,697
YOUNKIN/MAYO	706/835	616/925/0	1,541
MIRAGE	274/429	398/294/13	705
KRAMER/OHSU	142/188	59/109/162	330
ROSMAP	502/1,119	364/853/404	1,621
TGEN2	560/698	770/488/0	1,258
MIAMI/UMVUMSSM	817/1,380	1,085/1,112/0	2,197
KAMBOH2/UPITT	810/1,377	1,267/834/86	2,187
WASHU/GOATE	217/295	312/166/34	512
ACT 2	144/158	18/5/279	302
ADC 4	332/443	287/340/148	775
ADC 5	370/526	273/496/127	896
ADC 6	432/571	363/304/336	1,003
BIOCARD	75/113	8/123/57	188
CHAP	236/348	20/164/400	584
EAS	116/132	10/209/29	248
MTV	177/261	241/194/3	438
NBB	96/204	215/85/0	300
RMAYO	220/133	12/271/70	353
ROSMAP 2	105/323	62/237/129	428
TARC1	170/260	286/144/1	431
UKS	845/895	767/973/0	1,740
WASHU 2	68/67	30/65/40	135
WHICAP	246/394	74/562/4	640
Total	11,967/16,760	12,532/13,134/3,064	28,730

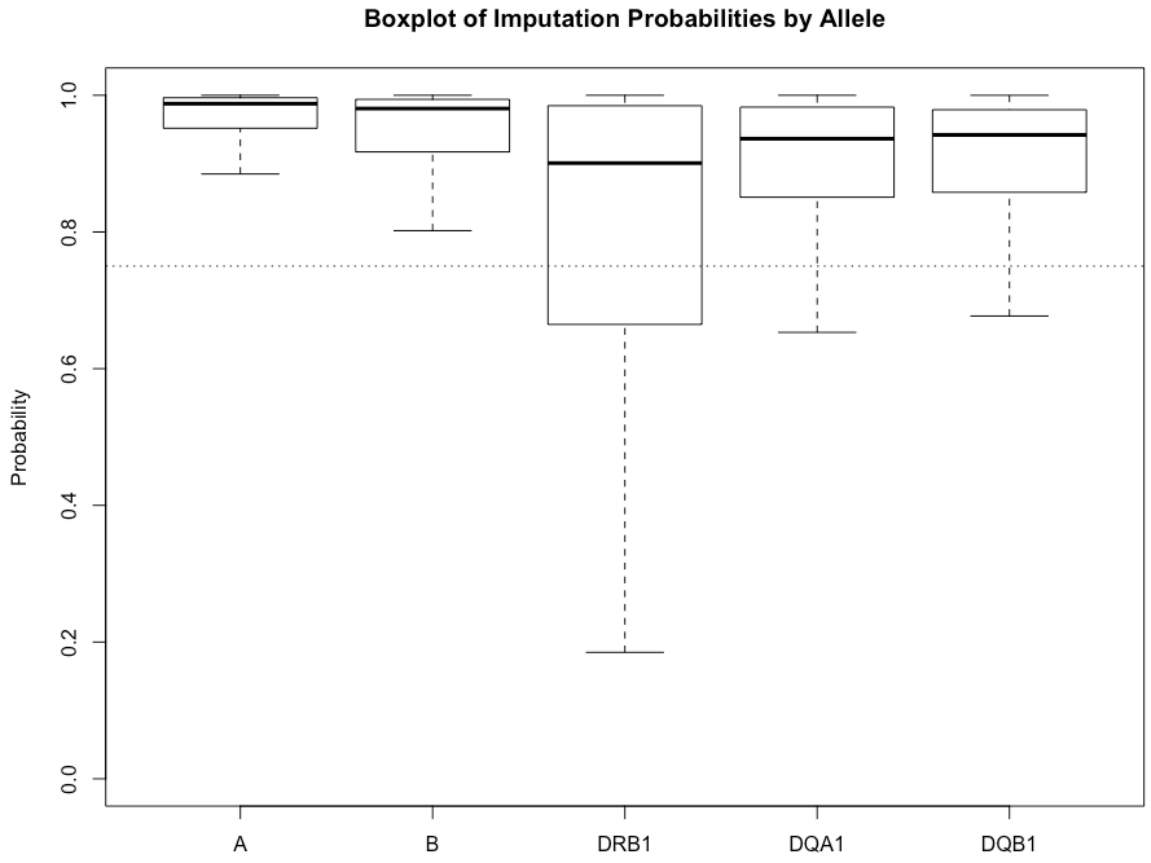
Study name, male/female distribution, case/control/missing distribution, and sample size of 30 datasets containing unrelated individuals combined into full ADGC dataset.

Table reproduced from:
Boehme KL, Mukherjee S, Crane PK, Kauwe JSK. ADGC 1000 genomes combined workflow (electronic document). September 2014. Accessible at: https://kauwelab.byu.edu/ADGC1KGImputation.aspx

ROSMAP References:
Bennett DA, Schneider JA, Arvanitakis Z, Wilson RS. Overview and findings from the Religious Orders Study. Current Alzheimer's Research. 2012;9:628-645. PMCID: PMC3409291

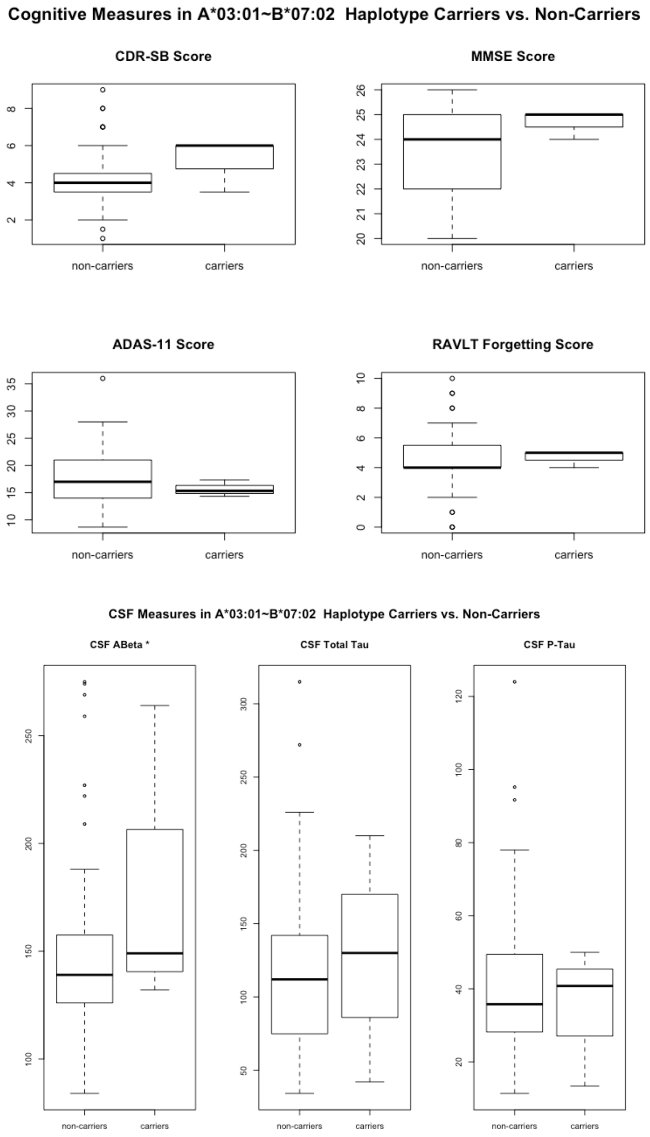
Bennett DA, Schneider JA, Buchman AS, Barnes LL, Boyle PA, Wilson RS. Overview and findings from the Rush Memory and Aging Project. Current Alzheimer's Research. 2012;9:646-663.

S1 Figure. Box-and-whisker plot of posterior probabilities of imputation for each of 5 imputed HLA-alleles in ADGC cohort.



Thick line represents median, box edges represent 1st and 3rd quartile, and whiskers represent 95% confidence interval. Values higher than 0.75 (dashed line) were included in the present study. Outlier dots are not shown for clarity.

S2 Figure. Box-and-whisker plots of AD cognitive and clinical biomarker measures in HLA *A*03:01~B*07:02* haplotype noncarriers and carriers in the ADNI cohort.



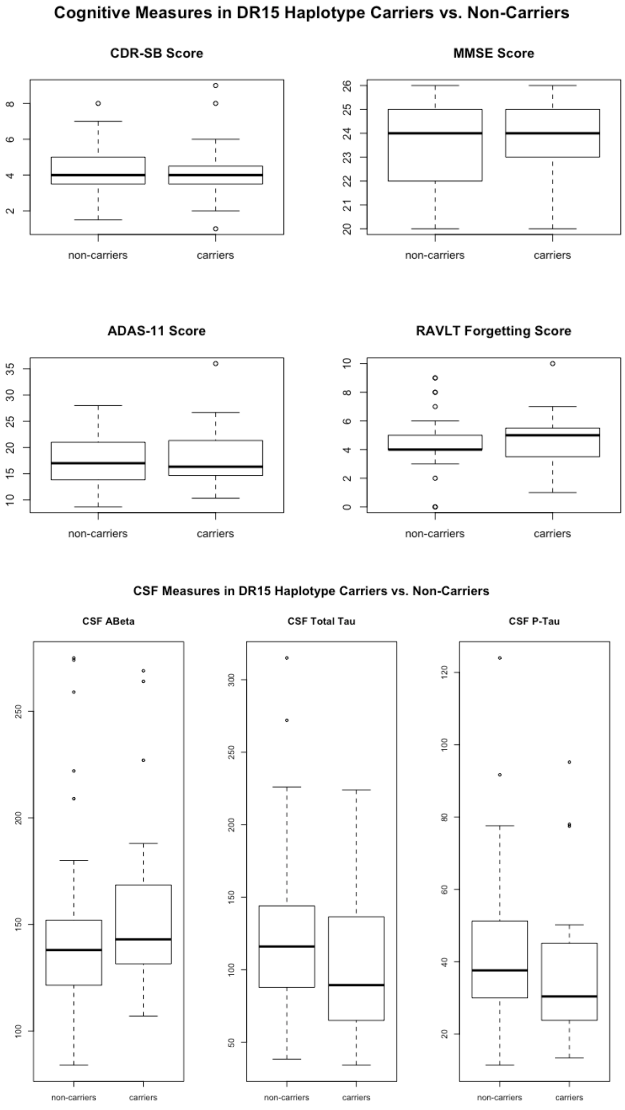
HLA A*03:01~B*07:02 haplotype carriers (n = 3) do not show any significant differences from haplotype non-carriers (n = 67) in a variety of cognitive assessments and measures of biomarkers in CSF in the ADNI cohort. The thick line represents the median, box edges represent the first and third quartiles, and whiskers represent the 95% CI. MMSE, Mini Mental State Exam; p-tau, phosphorylated tau.

S2 Table. Summary statistics for ADNI participants with baseline CSF protein measurements.

	CN	MCI	AD	*P*-Value
N	49	61	53	
Age (years)	75.6 ± 5.41	74.3 ± 6.59	75.8 ± 6.56	NS
Female (%)	57.1%	29.5%	47.2%	< 0.05
Education (years)	15.5 ± 2.54	15.3 ± 3.29	14.6 ± 3.12	NS
CDR-SB	0.14 ± 0.14	1.66 ± 0.89	4.18 ± 1.67	< 0.001
APOE-ε4 Carrier (%)	30.6%	60.7%	71.7%	< 0.001
Haplotype Dose (# single / # double)	20/1	15/0	20/1	NS
Chemokine CC-4 (ng/mL)	-1.47 ± 0.17	-1.41 ± 0.20	-1.49 ± 0.20	NS
ADAS (baseline)	5.86 ± 2.78	12.3 ± 3.41	17.9 ± 5.49	< 0.001
RAVLT Forgetting (baseline)	4.08 ± 3.02	5.13 ± 2.31	4.28 ± 2.03	NS

Descriptive data are summarized by diagnostic category. Values represent the mean ± standard error, percent, or number of participants in a given diagnostic category. Two-tailed p-values were from ANOVA (continuous traits) or chi-square (categorical values) tests by diagnostic group. CN, cognitively normal; NS, not significant ($p > 0.05$).

S3 Figure. Box-and-whisker plots of AD cognitive and clinical biomarker measures in *DR15* haplotype noncarriers and carriers in the ADNI cohort.



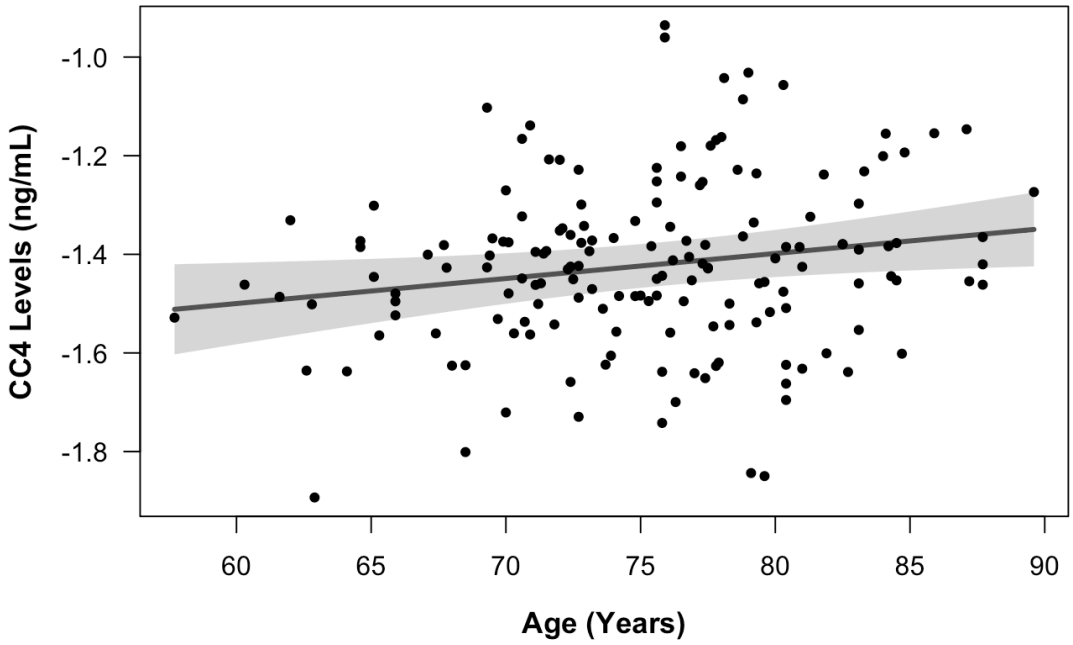
DR15 haplotype carriers (n = 23) do not show any significant differences from haplotype noncarriers (n = 47) in a variety of cognitive assessments and measures of biomarkers in CSF in the ADNI cohort. The thick line represents the median, box edges represent the first and third quartiles, and whiskers represent the 95% CI. MMSE, Mini Mental State Exam; p-tau, phosphorylated tau.

S3 Table. HLA *A*03:01~B*07:02* haplotype carriers do not show any significant baseline differences on clinical biomarker measures of AD with the exception of CSF amyloid β level.

Measurement	*P*-value
Volumetrics	
Whole brain volume	0.34
Ventricle volume	0.86
Middle temporal lobe volume	0.93
Hippocampus volume	0.86
Entorhinal cortex volume	0.42
Fusiform gyrus volume	0.41
Clinical measures	
CDR-SB score	0.38
MMSE score	0.32
ADAS11 score	0.52
RAVLT Forgetting score	0.97
Biomarker measures	
CSF amyloid beta level	0.03*
CSF total tau level	0.88
CSF p-tau level	0.74

Analysis of patients with baseline AD diagnosis from the ADNI cohort carrying HLA A*03:01~B*07:02 (n = 3) versus haplotype noncarriers (n = 63) shows no significant differences in volumetric, clinical, cognitive, and biomarker assessments relevant to AD with the exception of amyloid β levels as measured in CSF (p = 0.03). The p-values for volumetric measurements are the effect of carrying the DR15 haplotype (binary 0/1) in a linear regression model adjusted for baseline age, sex, years of education, dose of *APOE-ε4* allele (0/1/2), and intracranial volume. The p-values for clinical and biomarker measures are the effect of carrying the DR15 haplotype (binary 0/1) in a linear regression model adjusted for baseline age, sex, years of education, and dose of *APOE-ε4* allele (0/1/2). ADAS11, 11-item Alzheimer's Disease Assessment Scale cognitive subscale; MMSE, Mini-Mental State Exam.

S4 Figure. Chemokine CC4 levels are higher in older adults.



Chemokine CC4 levels are on average higher with greater age, suggesting higher levels of inflammation in older individuals. Chemokine CC4 levels are quality controlled and transformed as described in S1 Methods. The plotted points are partial residuals with 95% confidence bands provided in shading.

S4 Table. HLA *DR15* risk haplotype dosage is associated with longitudinal changes in ADAS and RAVLT cognitive test scores.

Outcome	Variable	Estimate ± SE	P-Value
ADAS (11 Item)	Age	0.02 ± 0.04	0.56
	Time	5.59 ± 2.40	0.01
	Sex	-0.68 ± 0.49	0.16
	CDR-SB	2.56 ± 0.12	$< 1x10^{-6}$
	Education	-0.16 ± 0.09	0.07
	APOE-ε4 Status	0.37 ± 0.36	0.31
	Haplotype Dose	-0.05 ± 0.47	0.92
	Time x Age	-0.07 ± 0.03	0.02
	Time x Sex	-0.07 ± 0.33	0.84
	Time x CDR-SB	0.82 ± 0.09	$< 1x10^{-6}$
	Time x Education	$-2.50 \times 10^{-3} \pm 0.06$	0.97
	Time x *APOE-ε4*	0.79 ± 0.25	1.20×10^{-3}
	Time x Haplotype Dose	**0.69 ± 0.32**	**0.03**
RAVLT Forgetting Score	Age	$8.89 \times 10^{-3} \pm 0.02$	0.6
	Time	0.45 ± 0.50	0.36
	Sex	0.25 ± 0.20	0.22
	CDR-SB	0.14 ± 0.05	0.01
	Education	$5.65 \times 10^{-3} \pm 0.04$	0.87
	APOE-ε4 Status	0.63 ± 0.15	1.00×10^{-4}
	Haplotype Dose	0.02 ± 0.20	0.91
	Time x Age	$9.70 \times 10^{-4} \pm 5.74 \times 10^{-3}$	0.87
	Time x Sex	-0.08 ± 0.06	0.22
	Time x CDR-SB	-0.08 ± 0.03	2.80×10^{-3}
	Time x Education	-0.02 ± 0.01	0.08
	Time x *APOE-ε4*	-0.10 ± 0.05	0.05
	Time x Haplotype Dose	**-0.15 ± 0.06**	**0.02**

Results from regression models used to determine the effect of dose of HLA risk haplotype *DRB1*15:01~DQA1*01:02~DQB1*06:02* on longitudinal changes in ADAS and RAVLT cognitive test scores in cognitively normal, MCI, and AD groups from the ADNI cohort. The beta estimate (Estimate) and accompanying standard error (SE) reflect the adjusted effect of each independent variable as a predictor of ADAS 11-item score and RAVLT forgetting index scores. HLA haplotype dose demonstrated a significant positive and negative association with the rate of change in the ADAS ($p = 0.02$) and RAVLT ($p = 0.03$) scores, respectively, across all diagnostic groups (Time x Haplotype Dose). In other words, a greater dose of risk haplotype was associated with worse decline in cognitive performance over time. For all disease groups, the linear statistical model included the following as independent variables: age, time (from baseline), sex, CDR-SB score, *APOE-ε4* carrier status, education, and haplotype dose. All tests were two-tailed.

S5 Table. *DR15* haplotype carriers do not show any significant baseline differences on clinical biomarker measures of AD.

Measurement	*P*-value
Volumetrics	
Whole brain volume	0.53
Ventricle volume	0.93
Middle temporal lobe volume	0.66
Hippocampus volume	0.51
Entorhinal cortex volume	0.56
Fusiform gyrus volume	0.57
Clinical measures	
CDR-SB score	0.92
MMSE score	0.58
ADAS11 score	0.51
RAVLT Forgetting score	0.86
Biomarker measures	
CSF amyloid beta level	0.46
CSF total tau level	0.27
CSF p-tau level	0.30

Analysis of patients with baseline AD diagnosis from the ADNI cohort carrying DR15 (n = 23) versus DR15 noncarriers (n = 47) show no significant differences in volumetric, clinical, cognitive, and biomarker assessments relevant to AD. The p-values for volumetric measurements are the effect of carrying the DR15 haplotype (binary 0/1) in a linear regression model adjusted for baseline age, sex, years of education, dose of *APOE-ε4* allele (0/1/2), and intracranial volume. The p-values for clinical and biomarker measures are the effect of carrying the DR15 haplotype (binary 0/1) in a linear regression model adjusted for baseline age, sex, years of education, and dose of *APOE-ε4* allele (0/1/2). MMSE, Mini-Mental State Exam.

S6 Table. HLA *DR15* risk haplotype dosage is associated with baseline levels of chemokine CC-4 in CSF.

Outcome	Variable	Estimate ± SE	P-Value
Chemokine CC-4	Age	$5.08 \times 10^{-3} \pm 2.23 \times 10^{-3}$	0.02
	Sex	-0.15 ± 0.03	5.84×10^{-7}
	CDR-SB	$-8.29 \times 10^{-3} \pm 7.09 \times 10^{-3}$	0.24
	Education	$-9.8 \times 10^{-3} \pm 4.63 \times 10^{-3}$	0.04
	APOE-ε4 Status	-0.01 ± 0.02	0.53
	Haplotype Dose	**0.08 ± 0.03**	**5.18×10^{-3}**

Regression models were used to determine the effect of HLA risk haplotype DRB1*15:01~DQA1*01:02~DQB1*06:02 dose on cross-sectional CSF levels of chemokine CC-4 in the cognitively normal, MCI, and AD groups of the ADNI cohort are summarized. The beta estimate (Estimate) and accompanying standard error (SE) reflect the adjusted effect of each independent variable as a predictor of chemokine CC-4 levels. HLA risk haplotype dose demonstrated a significant positive association with baseline chemokine CC-4 CSF levels across all diagnostic groups (Haplotype Dose, p = 5.18 x 10-3) such that more copies of the risk haplotype were associated with higher levels of CSF chemokine CC-4, a measure of inflammation. For all disease groups, the linear statistical model included the following as independent variables: age, sex, CDR-SB score, *APOE-ε4* carrier status, education, and haplotype dose. All tests were two-tailed.

Chapter III:

Direct sequencing of the *HLA* region identifies differential effect of risk haplotype in amnestic versus atypical Alzheimer's disease

Introduction

The role of the Human Leukocyte Antigen (HLA) region in neurodegenerative disease is becoming increasingly appreciated as numerous genome-wide association studies (GWAS) identify significant associations with this complex genomic region and disease risk. We previously used a robust imputation method on two case–control cohorts to identify HLA haplotypes associated with Alzheimer's disease (AD); this study is described in Chapter II [1]. Our cohorts consisted of a discovery cohort of 309 individuals from the UCSF Memory and Aging Center (MAC) and a replication cohort of 11,381 individuals from the Alzheimer's Disease Genetics Consortium (ADGC). In our UCSF discovery cohort, the haplotype *A*02:01~B*07:02~DRB1*15:01~DQA1*01:02 ~DQB1*06:02* was associated with AD risk (p = 0.01, odds ratio (OR) [95% confidence interval (CI)] = 3.69 [1.16–13.69]). In our replication cohort, a similar haplotype (identical at 4 of 5 loci), *A*03:01~B*07:02 ~DRB1*15:01~DQA1*01:02~DQB1*06:02*, was associated with AD risk (p = 8.5 x 10^{-4}, OR = 1.22 [1.08–1.38]). Separate analyses in the combined UCSF + ADGC cohort (n = 11,690) of class I and II haplotypes further supported the role of class I haplotype *A*03:01~B*07:02* (p = 0.03, OR = 1.11 [1.01–1.23]) and class II haplotype *DRB1*15:01~DQA1*01:02~DQB1*06:02* (*DR15*) (p = 0.03, OR = 1.08 [1.01–1.15]) as risk factors for AD. We followed up these findings to assess their relevance to disease in the Alzheimer's Disease Neuroimaging Initiative (ADNI) clinical dataset representing the spectrum of cognitively normal (CN) controls, individuals with mild cognitive impairment (MCI), and individuals with AD. Carrying the class I haplotype *A*03:01~B*07:02* was associated with higher CSF amyloid levels. We also found a dose-dependent association between the class II *DR15* haplotype and

greater rates of cognitive decline as measured by two cognitive assessments (ADAS-cog and RAVLT). In a subset of the same cohort, dose of *DR15* was also associated with higher baseline levels of chemokine CC-4, a biomarker of inflammation.

One of the main limitations of our previous study was the use of imputed genotypes and the availability of only a subset of HLA genes. While increasing evidence supports the high rate of concordance between imputed HLA alleles with sequencing results [2], the HLA region is genetically complex, with over 15,000 identified allelic variants [3]. In particular, imputation is critically dependent on an individual's self-reported ethnicity to ensure use of the most appropriate reference population, introducing the possibility for incorrect imputations based on a participant's self-reported information. We aimed to validate the genotypes imputed in our previous study through direct sequencing of the HLA region, as well as expand our analyses to additional HLA genes that we were not able to impute from SNP data (particularly HLA-*C* and –*DPB1*).

Along with confirming the validity of imputed genotypes, we also aimed to confirm the associations identified previously between HLA haplotypes and AD using directly sequenced genotypes in an expanded cohort. Finally, through sequencing of additional individuals with diagnoses beyond typical amnestic AD, we aimed to expand our analyses to compare and contrast HLA associations with amnestic AD versus atypical clinical forms, and to explore HLA risk contributions to frontotemporal dementia (FTD) spectrum disorders.

Materials and Methods

Participants

UCSF Memory and Aging Center (MAC) cohort: The research participants included in this study were 745 self-described white individuals over the age of 50 years. Individuals were clinically assessed at the UCSF MAC and were diagnosed through consensus conference with the following clinical syndromes, described in detail below:

- AD (n = 210)
- atypical AD: including AD (frontal), AD (language), posterior cortical atrophy (PCA), logopenic variant primary progressive aphasia (lvPPA), and AD with concomitant dementia with Lewy Bodies (DLB), progressive supranuclear palsy (PSP), or vascular disease (n = 118)
- behavioral variant of frontotemporal dementia (bvFTD; n = 35)
- semantic variant of primary progressive aphasia (svPPA): including predominantly left, predominantly right or unspecified svPPA (n = 66)
- clinically normal older adults recruited for studies of healthy aging (CN; n = 316)

308 individuals originally imputed for HLA genotypes were also sequenced to confirm imputation results (one individual from the original study (n = 309) did not have DNA available for sequencing). In addition, 437 samples were only sequenced. Individuals were seen at the UCSF MAC between 1999–2012 and were genotyped as part of their participation in longitudinal research on neurodegenerative disease and healthy cognitive aging. DNA was collected from 2000–2012, and genotyping (for participants with imputed genotypes) was performed in 2012. Because individuals are followed up

longitudinally, we verified clinical diagnosis at the beginning of a previous study (May 2015). A multidisciplinary team of neurologists, neuropsychologists, and nurses performed a detailed evaluation on all individuals and established a diagnosis according to consensus criteria for AD [4], atypical AD (which included AD (frontal) [4], AD (language) [4], PCA [4,5], lvPPA [4,6], AD with DLB [4,7], AD with PSP, or AD with vascular disease [4,8]), bvFTD [9], or svPPA [9]. Individuals included as controls underwent a similar assessment and were diagnosed as having normal cognition for their age. Participants who carried a known pathogenic variant for autosomal dominant neurodegenerative disease (*APP, PSEN1*, *PSEN2, GRN, MAPT, C9ORF72, TARDBP, FUS)* were excluded from this study. Participants or surrogates completed written informed consent for all genetic research related to neurodegenerative disease and healthy cognitive aging during their initial visit in accordance with the Institutional Review Board at the University of California, San Francisco, and the review board approved all aspects of this study.

Statistical analysis

Cohort demographic summary statistics. Summary statistics for participants' sex, age of onset, and *APOE-ε4* carrier status were calculated using R.

Imputation of HLA alleles. HLA genotypes were derived from chromosome 6 SNP data using an imputation program, HLA Genotype Imputation with Attribute Bagging (HIBAG) v1.3 [10], as described previously [1].

Sequencing of HLA alleles. HLA genotypes were directly sequenced from genomic DNA samples extracted using standard procedures and banked at the National Cell Repository for AD (NCRAD) or the UCSF MAC Neurosciences Clinical Research Unit.

High-resolution HLA typing was carried out at the American Society for Histocompatibility and Immunogenetics (ASHI)-accredited laboratory of Histogenetics, Inc (Ossining, NY). 3x-high resolution typing was carried out with Illumina technology on the MiSeq sequencing platform (Illumina Inc, San Diego, CA). A total 14 Class I and 8 Class II amplicons were used for 3x-high resolution typing. All Class I amplicons (for class I genes *HLA-A, -B, -C*) were generated by gene specific primers. Class II amplicons (for class II genes *HLA-DRB1, -DRB3*, *-DRB4*, *-DRB5*, *-DQB1* and *-DPB1)* were generated by gene specific and group specific primers. Each primer set was designed to anneal to introns flanking each desired exon. One bridge primer set that expands from exon 2 to exon 3 for each HLA-*A*, *-B* and *–C* gene was used to resolve haplotype ambiguities. Primer maps were used as described previously [11]. Alleles were provided in four-digit form.

Assessing imputed versus sequenced genotype concordance. All genotypes were given in four-digit form. Genotypes for the same gene and individual derived from imputation versus sequencing were classified as a full match if all four digits of the sequenced allele matched all four digits of the imputed allele. For example, an imputed genotype of *A*02:01* and a sequenced genotype of *A*02:01* would be considered a full match. Alleles were classified as a partial match if the first two digits of the sequenced allele matched the first two digits of the imputed allele. For example, an imputed genotype of *A*02:01* and a sequenced genotype of *A*02:02* would be considered a partial match. Alleles were classified as a mismatch if the first two digits of the sequenced allele did not match the first two digits of the imputed allele. For example, an imputed genotype of *A*02:01* and a sequenced genotype of *A*03:02* would be

considered a mismatch. Similarly, a sequenced genotype of *A*03:01* would be considered a mismatch based on first two digits not corresponding.

Calculating HLA associations with disease risk. OR estimates for participants with neurodegenerative diseases versus cognitively normal controls were calculated using a statistical package designed to specifically probe associations with the HLA (BIGDAWG, [12]), as described previously [1]. The phenotypes we focused on for this study were amnestic AD, atypical AD (including frontal/executive predominant, language predominant, PCA, lvPPA, and AD with concomitant vascular disease or DLB or PSP), bvFTD, and svPPA.

Results

HLA genotypes derived from sequencing largely agree with Imputed genotypes

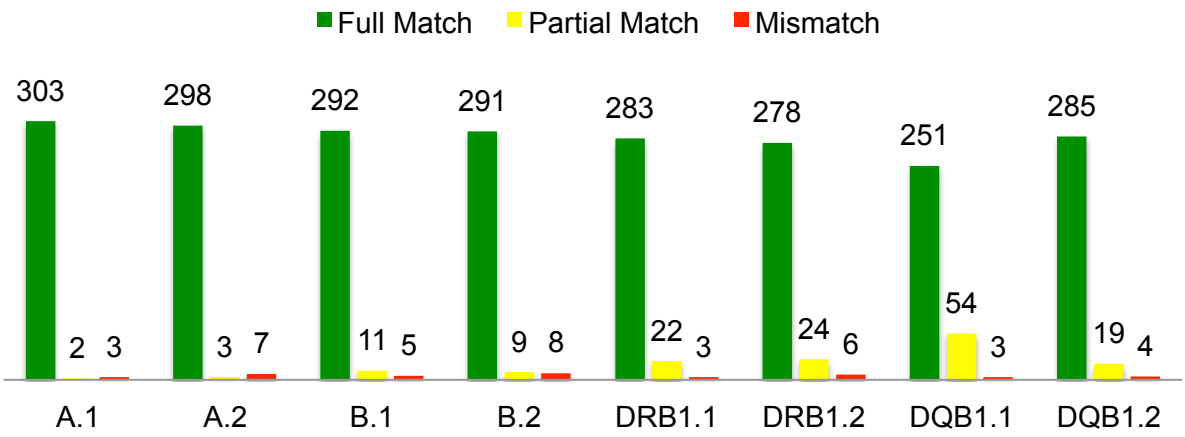
We directly sequenced nine HLA alleles: the class I alleles *A*, *B*, *C* and the class II alleles *DRB1*, *DRB3*, *DRB4*, *DRB5*, *DQB1* and *DPB1*. 308 of our sequenced samples had previously had HLA genotypes imputed for prior studies. HLA genotypes were previously imputed for class I alleles *A* and *B* and class II alleles *DRB1*, *DQA1,* and *DQB1* due to constraints of the imputation program. We compared the results for imputed genotype and sequenced genotype, i.e., for class I alleles *A* and *B* and class II alleles *DRB1* and *DQB1*, for all participants who had both types of data.

First, we analyzed the number of alleles that were a full match, a partial match, or a mismatch between imputed alleles and sequenced alleles in all 308 participants who had both types of genotypes. Overall, the sequenced alleles were mostly full matches with the imputed genotypes (Figure 1A). For *HLA-A* genotypes, which had a high

average call threshold (CT) of 0.96 suggesting generally high confidence in imputed results, full matches occurred in 97.6% of all alleles (Figure 1B). *HLA-B* genotypes were slightly less consistent than *A* but were still full matches in 94.6% of alleles (Figure 1B). *HLA-DRB1* genotypes had the lowest CT of all the compared alleles due to *DRB1*'s highly polymorphic nature, with a CT of 0.86 (Figure 1B). Overall, however, the *DRB1* sequenced genotypes matched relatively well with imputed genotypes, with 91.1% of alleles matching to four-digit resolution (full match) and an additional 7.5% of alleles matching at two-digit resolution (partial match). Less than 1.5% of *DRB1* alleles did not match between imputed and sequenced genotype (mismatch). *DQB1*, while expected to have a higher concordance between imputed and sequenced genotype than *DRB1* given its higher average CT value of 0.96, was the gene that showed the lowest rate of full four-digit matches in this study. While four-digit matches occurred in 87.0% of alleles, most of the alleles which did not match to four-digits matched to at least two-digits (11.9%) and only 1.1% of alleles were a complete mismatch (Figure 1B). Overall, of the 2,464 alleles compared in the full cohort (308 samples x 8 alleles), 2,281 (92.57%) were a full match, 144 (5.84%) were a partial match, and 39 (1.58%) were a mismatch (Figure 1C).

Figure 1. Comparison of imputed and sequenced genotypes in all Individuals.

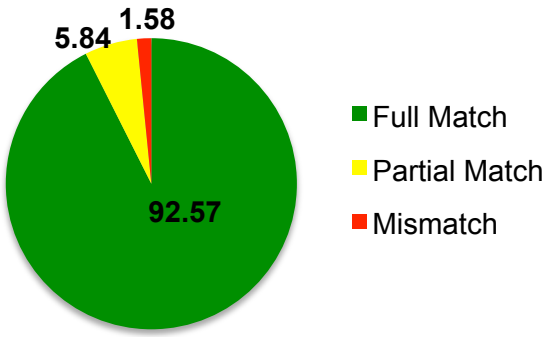
A. Number of total allele matches for each allele in full cohort.



B. Percentage of total allele matches for each allele in full cohort.

	A.1	A.2	B.1	B.2	DRB1.1	DRB1.2	DQB1.1	DQB1.2
Full Match	98.38	96.75	94.81	94.48	91.88	90.26	81.49	92.53
Partial Match	0.65	0.97	3.57	2.92	7.14	7.79	17.53	6.17
Mismatch	0.97	2.27	1.62	2.60	0.97	1.95	0.97	1.30
	0.96		0.91		0.86		0.96	
Average CT For Each Allele								

C. Percentage of total allele matches in full cohort.



Comparison of imputed and sequenced genotypes in all individuals with both data types available. A. Number of total allele matches (out of 308 samples) for each allele in full cohort. Full match denotes four-digit match, partial match denotes only two-digit match, and mismatch denotes no match. B. Percentage of allele matches in full cohort. Average CT for each allele is also shown, which provides information on the average confidence in imputation for each allele on a scale of 0 to 1 (1 being full confidence). C. Percentage of the total allele matches in the full cohort (out of 2,464 alleles total).

Sequencing of HLA alleles further validates quality threshold of CT > 0.75

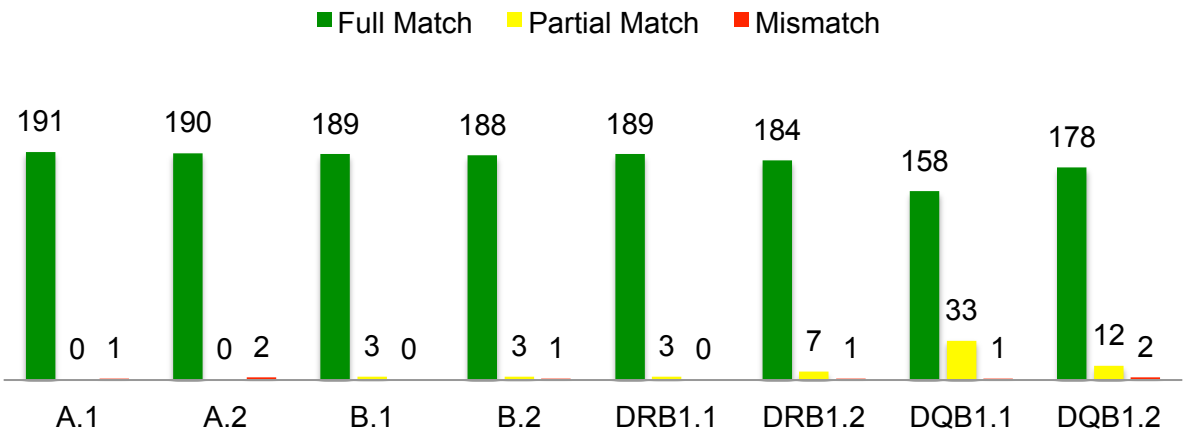
We repeated the analyses described above in a smaller cohort of individuals for whom we were more confident in imputation quality. This quality control (QC) threshold was used in our prior study [1] for imputed genotypes in the ADGC cohort (n = 11,381) and consisted of CT > 0.75 for all alleles for a given individual (for example, if a participant had CTs > 0.99 for *A*, *B*, and *DQB1* but had a CT of 0.74 for *DRB1*, they would not pass this threshold). This reduced our overall cohort size from 308 to 192, as 116 individuals had a CT $\leq$ 0.75 for at least one imputed genotype.

The implementation of our CT cutoff improved the concordance between imputed and sequenced genotypes (Figure 2). Average CT values for this cohort were expectedly higher based on our inclusion criteria; average CT values were 0.97-0.98 for all alleles in this cohort (Figure 2B). As in the full cohort, the *A* and *B* alleles showed very high concordance between imputed and sequenced genotypes, with > 98% of all *A* and *B* alleles matching to 4-digit resolution (Figure 2A, 2B). *DRB1* concordance was substantially increased with the implementation of the CT threshold, increasing four-digit match rate to 97.1%, with an additional 2.6% matching only at two-digit resolution, and only one allele (0.26% of all alleles compared) mismatched between imputed and sequenced genotypes (Figure 2B). Excluding CT values $\leq$ 0.75 did not have much effect on concordance between imputed and sequenced genotypes at *DQB1*, with 87.5% of alleles matching to four-digit resolution, 11.7% matching just at two-digit resolution, and 0.78% not matching (compared to 87.0%, 11.9%, and 1.1% in the full cohort prior to CT threshold implementation) (Figure 2B). Overall, of the 1,536 alleles compared in the cohort that passed our CT threshold (192 samples x 8 alleles), 1,497 (95.51%) were a

full match, 61 (3.97%) were a partial match, and only 8 (0.52%) were a mismatch (Figure 2C).

Figure 2. Comparison of imputed and sequenced genotypes in individuals passing CT threshold > 0.75 for all alleles.

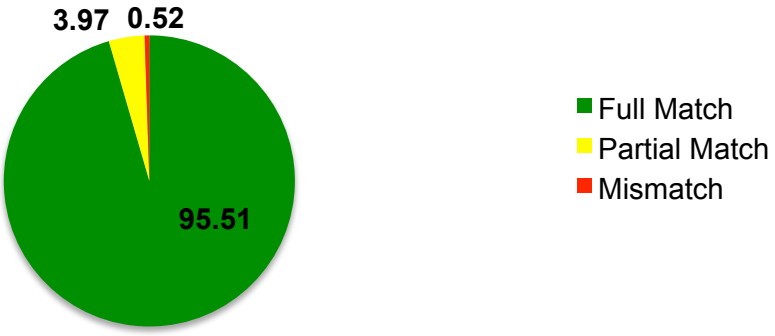
A. Number of total allele matches for each allele in cohort that passes CT threshold > 0.75.



B. Percentage of total allele matches for each allele in cohort that passes CT threshold > 0.75.

	A.1	A.2	B.1	B.2	DRB1.1	DRB1.2	DQB1.1	DQB1.2
Full Match	99.48	98.96	98.44	97.92	98.44	95.83	82.29	92.71
Partial Match	0.00	0.00	1.56	1.56	1.56	3.65	17.19	6.25
Mismatch	0.52	1.04	0.00	0.52	0.00	0.52	0.52	1.04
	0.98		0.97		0.97		0.98	
Average CT For Each Allele								

C. Percentage of total allele matches in cohort that passes CT threshold > 0.75.

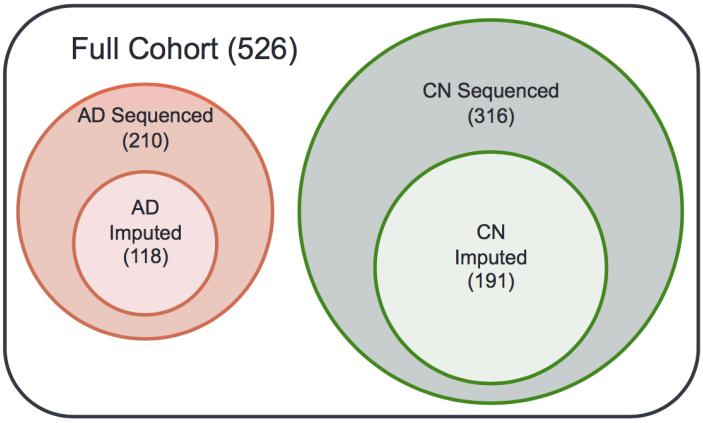


Comparison of imputed and sequenced genotypes in individuals passing CT threshold > 0.75 for all alleles. A. Number of total allele matches (out of 192 samples) for each allele in full cohort. Full match denotes four-digit match, partial match denotes only two-digit match, and mismatch denotes no match. B. Percentage of allele matches in full cohort and average CT for each allele. Average CT for each allele is also shown, which provides information on the average confidence in imputation for each allele on a scale of 0 to 1 (1 being full confidence). C. Percentage of the total allele matches in the full cohort (out of 1,536 alleles total).

Direct sequencing of HLA alleles confirms significant haplotypes and alleles associated with Alzheimer's disease risk

This cohort consisted of 526 individuals clinically diagnosed with AD (n = 210) or cognitively normal (CN) older adult controls (n = 316) with sequencing data for nine HLA alleles (HLA-*A*, -*B*, -*C*, -*DRB1*, -*DRB3*, -*DRB4*, -*DRB5*, -*DQB1* and -*DPB1*). Due to missing data in a large portion of participants at the *DRB3*, *DRB4* and *DRB5* alleles, we limited our analyses to the *A*, *B*, *C*, *DRB1*, *DQB1* and *DPB1* alleles, and haplotypes consisting of these alleles. Individuals missing sequenced alleles at any of these six alleles were excluded (n = 2 samples excluded). 308 of these individuals were included in our prior imputation study (see Chapter II) while 218 of these individuals had sequencing data only (no imputed genotypes) (Figure 3). Cohort summary statistics (for this typical AD cohort and all other diagnostic cohorts included in this study) are shown in Table 1.

Figure 3. Diagram of cohort distribution by diagnosis (AD vs. CN) and use in prior analysis with imputed genotypes.



Distribution of cohort by diagnosis (AD vs. CN) and use in prior analysis with imputed genotype (see Chapter II).

Table 1. Cohort demographics.

Cohort	*n*	% Male	*APOE-ε4* carriers (%)[#]	Age at onset[†]
AD (typical amnestic)	210	44.8	62.4	62.0 ± 9.6
atypical AD	118	55.9	48.3	60.5 ± 8.5
bvFTD	35	48.6	45.7	57.4 ± 19.4
svPPA	66	50.0	36.4	60.0 ± 6.9
cognitively normal (CN) controls	316	39.6	23.1	n/a

[#]Samples with available data: n = 209 amnestic AD, 118 atypical AD, 35 bvFTD, 66 svPPA, 316 CN
[†]Samples with available data: n = 97 amnestic AD, 44 atypical AD, 5 bvFTD, 31 svPPA, n/a for CN
Age at onset denotes mean ± standard deviation of age of first reported symptoms.

We performed association analysis on the four *A~B~C~DRB1~DQB1~DPB1* haplotypes with sufficient frequency in this small cohort. Of these four, one haplotype showed a significant association with AD risk: HLA *A*02:01~B*07:02~C*07:02~ DRB1*15:01~DQB1*06:02~DPB1*04:01* (*p* = 0.01, OR = 3.08 [1.15–9.08]) (Table 2). At the four alleles that overlap between this analysis and our prior analysis utilizing imputed genotypes (*A*, *B*, *DRB1* and *DQB1*), this haplotype is identical to the haplotype we previously identified as associated with AD risk in the UCSF cohort (n = 309): *A*02:01~B*07:02~DRB1*15:01~DQA1*01:02~DQB1*06:02* (*p* = 0.01, OR = 3.69 [1.16–13.69]) (see Chapter II, Table 2). This haplotype is also identical at three of four overlapping alleles to our most significant haplotype finding in the combined ADGC + UCSF cohort of 11,690 individuals: *A*03:01~B*07:02~DRB1*15:01~DQA1*01:02~ DQB1*06:02* (*p* = 9.6 x 10^{-4}, OR = 1.21 [1.08–1.37]) (see Chapter II, Table 3). Locus-level analyses of the present, sequenced cohort showed independent AD associations of *A*02:01*, *B*07:02*, *C*07:02*, *DRB1*15:01*, and *DQB1*06:02*, among other alleles (Table 3).

Table 2. Full haplotypes with significant risk associations with AD in combined cohort.

Full Haplotypes *A~B~C~DRB1~DQB1~DPB1*	OR (95% CI)	*p*-Value	CN freq	AD freq
01:01~08:01~07:01~03:01~02:01~01:01	*0.94 (0.24-3.28)*	*0.91 (NS)*	*0.0127*	*0.0119*
01:01~08:01~07:01~03:01~02:01~04:01	*0.77 (0.38-1.51)*	*0.42 (NS)*	*0.0459*	*0.0357*
02:01~**07:02**~07:02~**15:01**~**06:02**~04:01	3.08 (1.15-9.08)	0.01	0.0111	0.0333
*03:01~**07:02**~07:02~**15:01**~**06:02**~04:01*	*1.19 (0.62-2.2)*	*0.55 (NS)*	*0.0443*	*0.0523*

All analyzed haplotypes for directly sequenced cohort of AD vs. CN (*n* = 526). Non-significant associations (NS; p > 0.05) are shown in italics and grey text. Component alleles of the haplotype identified in our prior imputation study in the UCSF cohort are shown in bold. Component alleles of the haplotype identified in our prior imputation study in the ADGC cohort are shown underlined. In addition to OR with 95% CI, a breakdown of allele frequency in individuals with Alzheimer's disease (AD) versus cognitively normal (CN) older adult controls is provided.

Table 3. Individual alleles with significant risk associations with AD in sequenced cohort.

Class I Loci	OR (95% CI)	*p*-Value	CN frequency	AD frequency	European population frequency estimate
A*02:01	1.37 (1.02-1.83)	0.03	0.2184	0.2762	0.2960
B*07:02	1.75 (1.21-2.52)	1.8×10^{-3}	0.1108	0.1786	0.1400
B*35:02	0.25 (0.05-0.85)	0.02	0.0285	0.0071	0.0110
B*38:01	0.34 (0.13-0.80)	7.9×10^{-3}	0.0475	0.0167	0.0218
C*04:01	0.57 (0.37-0.87)	6.6×10^{-3}	0.1440	0.0881	0.1053
C*07:02	1.75 (1.22-2.51)	1.5×10^{-3}	0.1155	0.1857	0.1501
Class II Loci	**OR (95% CI)**	***p*-Value**	**CN frequency**	**AD frequency**	**European population frequency estimate**
DRB1*12:01	0.09 (0-0.60)	3.9×10^{-3}	0.0253	0.0024	0.0147
DRB1*15:01	1.44 (1.02-2.05)	0.03	0.1345	0.1833	0.1444
DQB1*03:01	0.66 (0.47-0.93)	0.02	0.2041	0.1452	0.1845
DQB1*06:02	1.41 (0.99-2.02)	0.048	0.1298	0.1738	0.1425

All significant loci results ($p < 0.05$) for directly sequenced cohort of AD vs. CN ($n = 526$). Component alleles of the haplotype identified in our prior imputation study in the UCSF cohort are shown in bold. Component alleles of the haplotype identified in our prior imputation study in the ADGC cohort are shown underlined. Component alleles of the most significant haplotype in this analysis (Table 2) are highlighted in blue. In addition to OR with 95% CI, a breakdown of allele frequency in individuals with Alzheimer's disease (AD) versus cognitively normal (CN) older adult controls is provided, in addition to the expected frequency in populations of European descent [13].

As in our prior analysis with imputed alleles, we also assessed class I (*HLA A~B~C*) and class II (*HLA DRB1~DQB1~DPB1*) haplotypes separately for their role in AD risk in this sequenced cohort. Out of 12 class I haplotypes available for analysis, two were significantly associated with AD ($p < 0.05$), including *A*02:01~B*07:02~C*07:02* ($p = 4.8 \times 10^{-3}$, OR = 2.63 [1.25-5.75]) (Table 4). Two class II haplotypes (out of 21 analyzed) were associated with AD risk (Table 4). The class II haplotype subset of our full six-allele haplotype finding (*DRB1*15:01~DQB1*06:02~DPB1*04:02*) was not significantly associated with AD, though it trended towards association with increased risk ($p = 0.07$, OR = 1.43 [0.95-2.16]).

Table 4. Separate class I and class II haplotypes with risk associations with AD in sequenced cohort.

Class I Haplotypes *A~B~C*	OR (95% CI)	*p*-Value	CN frequency	AD frequency
02:01~**<u>07:02</u>**~07:02	2.63 (1.25-5.75)	4.8×10^{-3}	0.0206	0.0524
26:01~38:01~12:03	0.18 (0.02-0.79)	0.01	0.0253	0.0048
Class II Haplotypes *DRB1~DQB1~DPB1*	**OR (95% CI)**	***p*-Value**	**CN frequency**	**AD frequency**
01:01~05:01~03:01	0.12 (0-0.84)	0.02	0.0190	0.0024
04:02~03:02~04:01	0.20 (0.02-0.85)	0.02	0.0237	0.0048

All significantly associated ($p < 0.05$) class I (*A~B~C*) and class II (*DRB1~DQB1~DPB1*) haplotypes for sequenced cohort of AD vs. CN ($n = 526$). Component alleles of the haplotype identified in our prior imputation study in the UCSF cohort are shown in bold. Component alleles of the haplotype identified in our prior imputation study in the ADGC cohort are shown underlined. Class I and class II component haplotypes of the most significant haplotype in this analysis (Table 2) are highlighted in blue. In total, 12 class I haplotypes and 21 class II haplotypes were analyzed. In addition to OR with 95% CI, a breakdown of allele frequency in individuals with Alzheimer's disease (AD) versus cognitively normal (CN) older adult controls is provided

HLA haplotype associated with amnestic AD risk shows protective effect in atypical AD patients

To assess whether HLA risk contributions are the same or different in individuals with atypical clinical presentations of AD, we performed HLA association analysis in patients diagnosed with atypical AD (including AD [frontal], AD [language], PCA, lvPPA, and AD with concomitant DLB, PSP or vascular disease [n = 118]) compared to cognitively normal controls (n = 316). Only one full six-allele haplotype was significantly associated with atypical AD: *A*03:01~B*07:02~C*07:02~ DRB1*15:01~DQB1*06:02~DPB1*04:01* (p = 0.01, OR = 0.18 [0.02-0.74]) (Table 5). This was the same haplotype identified in our previous analysis of 11,690 individuals as associated with risk for amnestic AD (Chapter II, Table 3), which overlaps at five of six alleles (all alleles but A) with our most significant finding for amnestic AD in the UCSF cohort (Table 2, also Chapter II, Table 2). Of note, this haplotype was associated with risk for amnestic AD but shows a protective effect in atypical AD. The imputed *A*03:01~B*07:02~DRB1*15:01 ~DQA1*01:02~DQB1*06:02* haplotype (the most similar haplotype we can assess, due to our inability to impute the *C* and *DPB1* alleles) identified in our previous analysis occurred at a frequency of 0.0472 in CN and 0.570 in typical AD in the ADGC cohort of over 11,000 individuals; the frequency of the sequenced *A*03:01~B*07:02~C*07:02 ~DRB1*15:01~DQB1*06:02~DPB1*04:01* haplotype in this cohort is 0.0443 in CN (similar to expected based on imputed study) and strikingly lower, only 0.0085, in atypical AD cases.

None of the component alleles of the full haplotype were significant at the locus level, though some other *C*, *DRB1*, *DQB1,* and *DPB1* alleles did show locus-level

significance (Table 6). Analysis of class I and class II genes separately showed that *A*03:01~B*07:02~C*07:02* was the only significant class I haplotype (of six analyzed) (p = 0.04, OR = 0.43 [0.16-0.98]) while *DRB1*15:01~DQB1*06:02~DPB1*04:01* was not significantly associated with disease (p = 0.36) (Table 7). Two other class II haplotypes were associated with either increased or decreased risk of atypical AD compared to controls (Table 7).

Table 5. Full haplotypes with significant risk associations with atypical AD in sequenced cohort.

Full Haplotypes *A~B~C~DRB1~DQB1~DPB1*	OR (95% CI)	*p*-Value	CN freq	Atypical AD freq
01:01~08:01~07:01~03:01~02:01~04:01	*0.59 (0.22-1.40)*	*0.21 (NS)*	*0.0491*	*0.0297*
03:01~**07:02**~07:02~**15:01**~**06:02**~04:01	0.18 (0.02-0.74)	0.01	0.0443	0.0085

All analyzed haplotypes for directly sequenced cohort of atypical AD vs. CN (*n* = 434). Non-significant associations (NS; p > 0.05) are shown in italics and grey text. Component alleles of the haplotype identified in our prior imputation study in the UCSF cohort are shown in bold. Component alleles of the haplotype identified in our prior imputation study in the ADGC cohort are shown underlined. In addition to OR with 95% CI, a breakdown of allele frequency in individuals with atypical Alzheimer's disease (AD) versus cognitively normal (CN) older adult controls is provided.

Table 6. Individual alleles with significant risk associations with atypical AD in sequenced cohort.

Class I Loci	OR (95% CI)	*p*-Value	CN frequency	Atypical AD frequency	European population frequency estimate
C*07:01	0.63 (0.39-0.98)	0.03	0.1883	0.1271	0.1666
C*08:02	1.82 (0.94-3.44)	0.047	0.0459	0.0805	0.0388
Class II Loci	**OR (95% CI)**	***p*-Value**	**CN frequency**	**Atypical AD frequency**	**European population frequency estimate**
DRB1*01:01	1.66 (0.99-2.74)	0.04	0.0807	0.1271	0.0915
DQB1*03:03	2.35 (1.11-4.91)	0.01	0.0301	0.0678	0.0446
DQB1*05:01	1.65 (1.08-2.50)	0.01	0.1250	0.1907	0.1223
DPB1*01:01	0.37 (0.13-0.88)	0.02	0.0665	0.0254	n/a

All significant loci results ($p < 0.05$) for directly sequenced cohort of atypical AD vs. CN ($n = 434$). In addition to OR with 95% CI, a breakdown of allele frequency in individuals with atypical Alzheimer's disease (AD) versus cognitively normal (CN) older adult controls is provided, in addition to the expected frequency in populations of European descent [13].

Table 7. Separate class I and class II haplotypes with significant risk associations with atypical AD in sequenced cohort.

Class I Haplotypes *A~B~C*	OR (95% CI)	*p*-Value	CN frequency	Atypical AD frequency
03:01~**07:02**~07:02	0.43 (0.16-0.98)	0.04	0.0665	0.0297
Class II Haplotypes *DRB1~DQB1~DPB1*	**OR (95% CI)**	***p*-Value**	**CN frequency**	**Atypical AD frequency**
07:01~03:03~04:01	2.53 (0.99-6.34)	0.02	0.0190	0.0466
11:01~03:01~04:01	0.23 (0.03-0.93)	0.03	0.0364	0.0085

All significantly associated ($p < 0.05$) class I (*A~B~C*) and class II (*DRB1~DQB1~DPB1*) haplotypes for sequenced cohort of atypical AD vs. CN ($n = 434$). Component alleles of the haplotype identified in our prior imputation study in the UCSF cohort are shown in bold. Component alleles of the haplotype identified in our prior imputation study in the ADGC cohort are shown underlined. In total, six class I haplotypes and eight class II haplotypes were analyzed. In addition to OR with 95% CI, a breakdown of allele frequency in individuals with atypical Alzheimer's disease (AD) versus cognitively normal (CN) older adult controls is provided.

Comparison between haplotypes associated with typical amnestic AD vs. atypical AD highlights differential effect of top haplotype on AD subtypes

We were intrigued by the observation that the effect of haplotypes including *B*07:02~DRB1*15:01~DQB1*06:02* in typical amnestic AD cases vs. controls (risk effect) was in the opposite direction in the comparison of atypical AD cases vs. controls (protective effect). In order to rule out the possibility that the control cohort was driving the results and to directly test the hypothesis that this haplotype has opposing risk effects in typical amnestic versus atypical AD, we directly tested for HLA haplotypes associated with atypical AD cases versus typical amnestic AD cases.

Of three full six-allele haplotypes available to test, the only significant association was with *A*03:01~B*07:02~C*07:02~DRB1*15:01~DQB1*06:02~DPB1*04:01*, (p = 4.0 x 10^{-3}, OR = 0.15 [0.02-0.64]) (Table 8), which was the same haplotype associated with decreased risk for atypical AD compared to controls (Table 5). This haplotype is identical at five of six loci (all loci except *A*) to the haplotype associated with greater risk for typical amnestic AD compared to controls (Table 2). At the locus level, four of the six component alleles of this haplotype were associated with decreased risk of atypical AD vs. amnestic AD: *B*07:02*, *C*07:02*, *DRB1*15:01,* and *DQB1*06:02* (Table 9). When analyzing class I and class II haplotypes separately in atypical AD vs. typical amnestic AD, the only significant class I finding was *A*02:01~B*07:02~C*07:02* (p = 0.03, OR = 0.36 [0.11-0.97]) and the only significant class II finding was *DRB1*15:01~DQB1*06:02 ~DPB1*04:01* (p = 0.02, OR = 0.52 [0.28-0.92]) (Table 10).

Table 8. Full haplotypes with significant risk associations in atypical vs. typical AD.

Full Haplotypes *A~B~C~DRB1~DQB1~DPB1*	OR (95% CI)	*p*-Value	Typical AD freq	Atypical AD freq
01:01~08:01~07:01~03:01~02:01~04:01	*0.72 (0.25-1.87)*	*0.48 (NS)*	*0.0405*	*0.0297*
***02:01*~*07:02*~07:02~*15:01*~*06:02*~04:01**	*0.54 (0.13-1.78)*	*0.28 (NS)*	*0.0310*	*0.0169*
03:01~**07:02**~07:02~**15:01**~**06:02**~04:01	0.15 (0.02-0.64)	4.0×10^{-3}	0.0524	0.0085

All analyzed haplotypes for sequenced cohort of atypical AD vs. AD (*n* = 328). Non-significant associations (NS; p > 0.05) are shown in italics and grey text. Component alleles of the haplotype identified in our prior imputation study in the UCSF cohort are shown in bold. Component alleles of the haplotype identified in our prior imputation study in the ADGC cohort are shown underlined. The haplotype identified in Table 2 is highlighted in blue. In addition to OR with 95% CI, a breakdown of allele frequency in individuals with atypical Alzheimer's disease (AD) versus AD is provided.

Table 9. Individual alleles with significant risk associations in atypical vs. typical AD.

Class I Loci	OR (95% CI)	*p*-Value	Typical AD frequency	Atypical AD frequency	European population frequency estimate
A*26:01	2.31 (0.99-5.50)	0.03	0.0286	0.0636	0.0295
B*07:02	0.43 (0.24-0.73)	1.0×10^{-3}	0.1786	0.0848	0.1399
B*38:01	2.61 (0.88-8.19)	0.047	0.0167	0.0424	0.0218
B*44:02	1.86 (0.96-3.61)	0.045	0.0524	0.0932	0.0901
C*07:02	0.47 (0.28-0.79)	2.6×10^{-3}	0.1857	0.0975	0.1501
C*08:02	2.21 (1.05-4.69)	0.02	0.0381	0.0805	0.0388
Class II Loci	**OR (95% CI)**	***p*-Value**	**Typical AD frequency**	**Atypical AD frequency**	**European population frequency estimate**
DRB1*01:01	2.63 (1.43-4.92)	6.7×10^{-4}	0.0524	0.1271	0.0915
DRB1*03:01	0.56 (0.32-0.94)	0.02	0.1619	0.0975	0.1292
DRB1*15:01	0.48 (0.28-0.80)	3.3×10^{-3}	0.1833	0.0975	0.1444
DQB1*02:01	0.63 (0.42-0.95)	0.02	0.2667	0.1864	0.2303
DQB1*05:01	2.24 (1.38-3.64)	4.8×10^{-4}	0.0952	0.1907	0.1228
DQB1*06:02	0.51 (0.30-0.86)	7.9×10^{-3}	0.1738	0.0975	0.1425
DPB1*01:01	0.41 (0.14-1.05)	0.048	0.0595	0.0254	n/a

All significant loci results ($p < 0.05$) for sequenced cohort of atypical AD vs. AD ($n = 328$). Component alleles of the haplotype identified in our prior imputation study in the UCSF cohort are shown in bold. Component alleles of the haplotype identified in our prior imputation study in the ADGC cohort are shown underlined. Component alleles of the haplotype identified in Table 2 are highlighted in blue. In addition to OR with 95% CI, a breakdown of allele frequency in individuals with atypical Alzheimer's disease (AD) versus AD is provided in addition to the expected frequency in populations of European descent [13].

Table 10. Separate class I and class II haplotypes with significant risk associations in atypical vs. typical AD.

Class I Haplotypes *A~B~C*	OR (95% CI)	*p*-Value	Typical AD frequency	Atypical AD frequency
02:01~**07:02**~07:02	0.36 (0.11-0.97)	0.03	0.0571	0.0212
Class II Haplotypes *DRB1~DQB1~DPB1*	**OR (95% CI)**	***p*-Value**	**Typical AD frequency**	**Atypical AD frequency**
15:01~**06:02**~04:01	0.52 (0.28-0.92)	0.02	0.1381	0.0763

All significantly associated ($p < 0.05$) class I (*A~B~C*) and class II (*DRB1~DQB1~DPB1*) haplotypes for sequenced cohort of atypical AD vs. AD ($n = 328$). Component alleles of the haplotype identified in our prior imputation study in the UCSF cohort are shown in bold. Component alleles of the haplotype identified in our prior imputation study in the ADGC cohort are shown underlined. Component alleles of the haplotype identified in Table 2 are highlighted in blue. In total, five class I haplotypes and seven class II haplotypes were analyzed. In addition to OR with 95% CI, a breakdown of allele frequency in individuals with atypical Alzheimer's disease (AD) versus AD is provided.

B*07:02~C*07:02~DRB1*15:01~DQB1*06:02~DPB1*04:01 effect by APOE-ε4 status and sex.

Based on our previous finding that the *A*03:01~B*07:02~DRB1*15:01~DQA1*01:02 ~DQB1*06:02* haplotype was primarily associated with AD risk in men but not women, and seemed to have a stronger risk effect in *APOE-ε4* non-carriers, we analyzed the effect of our top haplotype in cohorts subdivided by *APOE-ε4* carrier status or by sex (*ε4* carriers only, *ε4* non-carriers only; males only, females only). Given the unclear effects of variation at the *HLA-A* locus (i.e., *A*02:01* and *A*03:01*) in our previous analyses, we omitted the *HLA-A* allele from these subgroup analyses and only examined the *B*07:02~C*07:02~DRB1*15:01~DQB1*06:02~DPB1*04:01* haplotype. This allowed us to increase our power to detect significant between-group effects since the frequency of this five-allele haplotype was higher. Frequencies of the haplotype in each subgroup are shown in Table 11 (by *APOE-ε4* carrier status) and Table 12 (by sex).

Many of the subgroup analyses showed continued significant associations between disease risk and the *B*07:02~C*07:02~DRB1*15:01~DQB1*06:02 ~DPB1*04:01* haplotype. Three of the four subgroup analyses showed significant associations with the haplotype of interest in the atypical AD vs. AD comparison; the association was not statistically significant in *APOE-ε4* non-carriers but the *p*-value was very close to the threshold for significance (p = 0.055). The only association that showed a clear differential effect by subgrouping was the atypical AD vs. CN comparison, in which the haplotype had a significant protective effect in men but not in women (p = 0.025, OR = 0.27 [0.05-0.93] for men). Taken together, these findings suggest that the sex-specific effect of HLA-associated risk we observed in typical

amnestic AD may not be as robust in atypical forms of AD, though our limited sample sizes prevent us from drawing definitive conclusions on sex-specific effects of HLA loci on atypical AD risk.

Table 11. Frequencies of *B*07:02~C*07:02~DRB1*15:01~DQB1*06:02~DPB1*04:01* haplotype by *APOE-ε4* status.

A. AD vs. CN

	***ε4* non-carriers**		***ε4* carriers**	
Full Haplotypes *B~C~DRB1~DQB1~DPB1*	**CN freq**	**AD freq**	**CN freq**	**AD freq**
07:02~07:02~15:01~06:02~04:01	0.0700	0.0897	0.0616	0.1221

B. Atypical AD vs. CN

	***ε4* non-carriers**		***ε4* carriers**	
Full Haplotypes *B~C~DRB1~DQB1~DPB1*	**CN freq**	**Atypical AD freq**	**CN freq**	**Atypical AD freq**
07:02~07:02~15:01~06:02~04:01	0.0700	0.0328	0.0548	0.0351

C. Atypical AD vs. AD

	***ε4* non-carriers**		***ε4* carriers**	
Full Haplotypes *B~C~DRB1~DQB1~DPB1*	**AD freq**	**Atypical AD freq**	**AD freq**	**Atypical AD freq**
07:02~07:02~15:01~06:02~04:01	0.0897	0.0328	0.1183	0.0351

Frequency of the *B*07:02~C*07:02~DRB1*15:01~DQB1*06:02~DPB1*04:01* haplotype by sub-cohort divided by *APOE-ε4* carrier status and diagnosis.

Table 12. Frequencies of *B*07:02~C*07:02~DRB1*15:01~DQB1*06:02~DPB1*04:01* haplotype by sex.

A. AD vs. CN

	males		females	
Full Haplotypes *B~C~DRB1~DQB1~DPB1*	**CN freq**	**AD freq**	**CN freq**	**AD freq**
07:02~07:02~15:01~06:02~04:01	0.0800	0.1223	0.0602	0.1034

B. Atypical AD vs. CN

	males		females	
Full Haplotypes *B~C~DRB1~DQB1~DPB1*	**CN freq**	**Atypical AD freq**	**CN freq**	**Atypical AD freq**
07:02~07:02~15:01~06:02~04:01	0.0800	0.0227	0.0576	0.0385

C. Atypical AD vs. AD

	males		females	
Full Haplotypes *B~C~DRB1~DQB1~DPB1*	**AD freq**	**Atypical AD freq**	**AD freq**	**Atypical AD freq**
07:02~07:02~15:01~06:02~04:01	0.1223	0.0303	0.1034	0.0385

Frequency of the *B*07:02~C*07:02~DRB1*15:01~DQB1*06:02~DPB1*04:01* haplotype by sub-cohort divided by sex and diagnosis.

Small cohorts of bvFTD and svPPA patients suggest potential loci involved in disease risk

In addition to assessing the association between HLA haplotypes and alleles in AD (both typical and atypical), we also examined associations to small cohorts comprised of patients diagnosed with two other forms of neurodegenerative disease, bvFTD and svPPA. Due to modest statistical power from small sample sizes (n = 35 bvFTD patients, n = 66 svPPA patients), we did not find any haplotypes (either full six-allele or class I or class II subsets) associated with either of these diseases. However, locus level analysis did yield a handful of significant associations (all with class I alleles) with disease risk. Two loci were associated with bvFTD risk: *B*44:02* and *C*07:02* (Table 13). Interestingly, both of these alleles were significantly associated with AD in at least one of our previous analyses. *B*44:02* was significantly associated with greater atypical AD risk as compared to typical AD (Table 9). *C*07:02* was associated with greater risk of typical AD compared to controls (Table 3), and with decreased risk of atypical AD compared to typical AD (Table 9). One locus was significantly associated with svPPA risk: *C*04:01* (Table 14). *C*04:01* was also associated with decreased risk of typical AD compared to controls (Table 3).

Table 13. Individual alleles with significant risk associations in bvFTD.

Class I Loci	OR (95% CI)	*p*-Value	CN frequency	bvFTD frequency	European population frequency estimate
B*44:02	2.28 (0.97-4.91)	0.02	0.0680	0.1429	0.0901
C*07:02	2.09 (1.04-3.98)	0.02	0.1155	0.2143	0.1501

All significant loci results ($p < 0.05$) for sequenced cohort of bvFTD vs. CN ($n = 351$). Component alleles of the most significant haplotype in our AD association analysis (Table 1) are highlighted in blue. In addition to OR with 95% CI, a breakdown of allele frequency in individuals with bvFTD versus cognitively normal (CN) older adult controls is also provided in addition to the expected frequency in populations of European descent [13].

Table 14. Individual alleles with significant risk associations in svPPA.

Class I Loci	OR (95% CI)	*p*-Value	CN frequency	svPPA frequency	European population frequency estimate
C*04:01	0.49 (0.22-0.97)	0.04	0.1340	0.0758	0.1053

All significant loci results ($p < 0.05$) for sequenced cohort of svPPA vs. CN ($n = 382$). In addition to OR with 95% CI, a breakdown of allele frequency in individuals with svPPA versus cognitively normal (CN) older adult controls is also provided in addition to the expected frequency in populations of European descent [13].

Discussion

The first goal of our study was to validate imputed HLA genotypes through direct sequencing. In line with recent studies supporting the high rate of concordance between imputed HLA genotypes and direct sequencing (particularly in individuals of European descent) [2], our study validates HLA imputation as an effective method for assessing associations between HLA haplotypes and disease. Furthermore, the high percentage of concordance with the use of a CT threshold > 0.75 provides further confidence in the findings from the much larger ADGC cohort (in which the QC threshold was applied) used in our previous study, which we were not able to confirm by direct sequencing as we did with the UCSF MAC cohort.

Next, our study aimed to confirm HLA associations with AD previously identified utilizing imputed genotypes. Directly sequencing the HLA region in 308 of the 309 individuals used in this prior study, and adding an additional 218 individuals, confirmed the association of the *A*02:01~B*07:02~C*07:02~DRB1*15:01~DQB1*06:02~ DPB1*04:01* haplotype with AD, with a similar effect size, confidence interval, and p-value to what we previously found (p = 0.01, OR = 3.08 [1.15–9.08] in this study). In our original study, the haplotype frequency of a similar haplotype (*A*02:01~B*07:02 ~DRB1*15:01~DQA1*01:02~DQB1*06:02*) was 0.0131 in controls, versus 0.0466 in AD patients did. Due to the addition of 218 new samples, we found slightly different frequencies of this haplotype in our expanded, sequenced cohort. In this study, controls showed a similar haplotype frequency (0.0111) while a somewhat smaller frequency was seen in AD patients (0.0333) did; we suspect this is due to the inclusion of the

additional *C*07:02* and *DPB1*04:01* alleles, which are frequently but not always in linkage disequilibrium (LD) with the other component alleles of this haplotype.

Lastly, our study aimed to expand HLA association analysis to probe differences between amnestic versus atypical AD and to explore risk contributions to FTD spectrum disorders. The only significantly associated haplotype identified for atypical AD was *A*03:01~B*07:02~C*07:02~DRB1*15:01~DQB1*06:02~DPB1*04:01*, which we have previously found to be associated with increased risk for 'typical' amnestic AD in a large cohort of 11,690 individuals. Intriguingly, this haplotype was associated with the opposite risk effect for atypical AD, with atypical AD patients showing lower frequencies of this 'risk' haplotype when compared to healthy controls. We further corroborated this finding by directly assessing HLA differences between typical and atypical AD groups and indeed found that the risk haplotype in amnestic AD patients occurs at a significantly lower frequency in atypical AD patients. We hypothesize that this haplotype modifies the age of onset and/or neuroanatomical vulnerability to AD pathology, which could explain why the frequency is so strikingly different in each patient group.

In general, the genetic contribution to atypical clinical forms of AD is less well understood than the genetics underlying 'typical' amnestic AD. Previous studies have shown that the *APOE-ε4* risk factor is less frequent in atypical AD cases, such as dysexecutive AD, compared to typical amnestic AD [14,15]. Visuospatially affected PCA patients have been shown across studies to have less or even no risk association with *APOE-ε4* [16,17]. As expected, atypical AD individuals in our cohort had a lower frequency of *APOE-ε4* alleles (48.3% *APOE-ε4* carriers vs. 62.4% *APOE-ε4* carriers in amnestic AD cohort, compared to 23.1% in controls). While the lower frequency of

APOE-ε4 carriers in atypical AD suggests other genetic factors may contribute to atypical forms of disease, few genetic variants have been identified to account for these differences. A small number of genetic variants have been linked with PCA (including variation in the immune-related genes *CR1*, *TREM2* and potentially *CLU* as well) [17–19], lvPPA [20] and other atypical AD diagnoses, with some variants showing similar effects in typical amnestic late-onset AD (LOAD) risk and other variants showing disparate effects (such as *APOE-ε4*).

Of particular relevance to our results, a recent study found that the rs9271192 SNP, proxying *HLA-DRB5/HLA-DRB1*, while associated with an increased OR in previous GWAS of AD [21], is associated with a protective effect in AD patients presenting with predominant deficits in language or executive functioning (OR = 0.65 for language-predominant, 0.55 for dysexecutive-predominant) [22]. While this study was small (n = 38 for language-predominant AD and n = 37 for executive-predominant AD) and only assessed one SNP tagging the HLA region (no information on specific alleles), these findings are consistent with our present study. Together, these results support the idea that the HLA region and, in particular, the *A*03:01~B*07:02~C*07:02 ~DRB1*15:01~DQB1*06:02~DPB1*04:01* haplotype may exert strikingly different effects in atypical presentations of AD compared with amnestic AD. While our sample size for atypical AD cases is limited, our results support increasing evidence that atypical AD presentations should be considered distinct biological entities from typical amnestic LOAD, with different underlying genetic risk contributions and neuroanatomical vulnerability despite shared protein pathology.

Given increasing evidence that the *HLA* region and specific component alleles within this haplotype are associated with risk for numerous neurodegenerative disorders, including multiple sclerosis [23–25], Parkinson's disease (PD) [26,27], and FTD [28], as well as other neurological disorders including narcolepsy [29,30] and autism [31,32], our results of differential risk effects on typical versus atypical AD suggest an increasingly complex role of HLA receptors in neurological and neurodegenerative diseases. Though striking, this finding is consistent with observations in other classes of disease. Within autoimmune disorders, it is well established that specific haplotypes/alleles have opposing risk effects in different diseases. For example, the class II haplotype *DR15* is known to increase risk for multiple sclerosis and systemic lupus erythematosus, yet *DR15* is associated with reduced risk of Type 1 diabetes, and its component allele *DRB1*15:01* is associated with decreased risk of rheumatoid arthritis [33].

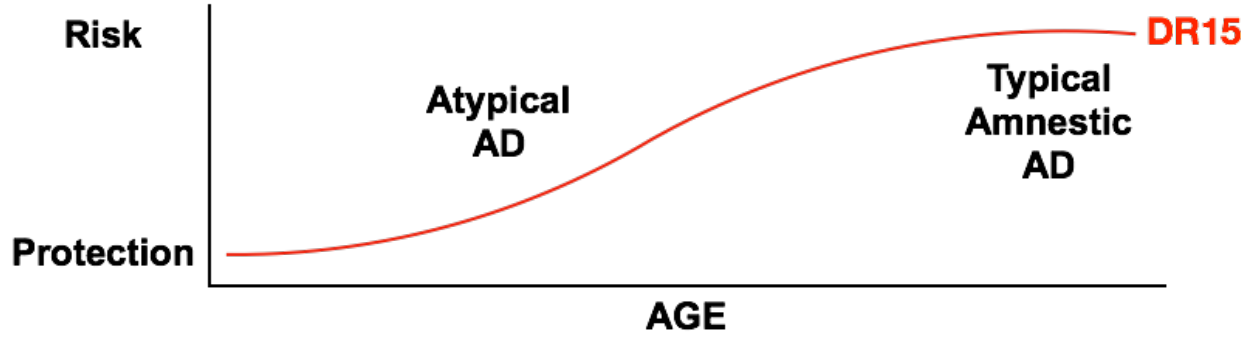
We have several hypotheses for how our identified HLA haplotype may exert dichotomous effects on amnestic versus atypical AD, which we will explore in future studies. Our results suggest that the *B*07:02~C*07:02~DRB1*15:01~DQB1*06:02 ~DPB1*04:01* haplotype exerts a disease modifying effect in AD, potentially through physical differences in the way the immune system is responding to Alzheimer's pathology, differences in HLA receptors across different brain regions, and/or through other indirect effects on neuroanatomical vulnerability. As atypical AD presentations are associated with lower *APOE-ε4* carrier frequencies, occur in higher proportions of males, and most often have an earlier age of onset compared to amnestic-predominant syndromes, we hypothesized that the identified haplotype may be interacting with these

demographic factors to modify disease risk in atypical versus amnestic AD. Our previous study of typical AD suggested that the *A*03:01~B*07:02~DRB1*15:01 ~DQA1*01:02~DQB1*06:02* haplotype was only a risk factor in individuals who did not carry an *APOE-ε4* allele, and was a stronger risk factor in men versus women [1]. However, we did not observe any disparate significant associations of the *B*07:02~C*07:02~DRB1*15:01~DQB1*06:02~DPB1*04:01* haplotype in subgroups divided by *APOE-ε4* carrier status or sex, with the exception of a significant association of the protective haplotype with atypical AD (compared to CN) in men only. One potential explanation for this is that the effect of the haplotype is dependent on age; at younger ages, when atypical AD is more likely to occur, this haplotype could be protective through the effects of its encoded receptors on 'good' inflammatory processes (Figure 4A). However as age increases, the effect of the same haplotype and its encoded receptors may transition from being helpful to harmful, perhaps through aberrant pro-inflammatory activities, thus increasing the risk of carriers developing typical amnestic AD. Presumably, the shift from being 'good' to 'bad' would be triggered by other, independent effects related to aging such as reduced capacity to clear protein aggregation or increased neuroinflammatory milieu. A second potential explanation for our observed differential HLA effect is that predisposition to neurodegeneration may be associated with developmental factors; for example, patients with lvPPA have higher frequencies of learning disabilities earlier in life [34]. It is possible that these neuroanatomical vulnerabilities interact with HLA-associated effects to alter predisposition to clinical manifestation of AD pathology in different phases of life. A third explanation is the possibility that the driving effect here is towards a risk of typical

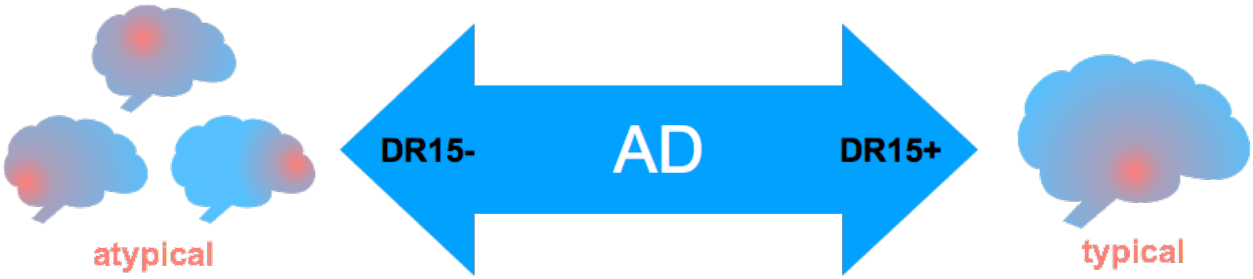
amnestic Alzheimer's, rather than a protective effect from atypical AD; other factors may promote AD pathological risk and the extremely low frequency of this haplotype in atypical AD cases may simply be due to a depletion effect, with carriers being more likely to develop typical AD versus atypical AD (Figure 4B). The specific mechanisms through which this haplotype exerts such disparate effects in typical versus atypical forms of AD remains to be elucidated in larger cohorts and biological studies designed to directly probe the effects of HLA receptors on AD pathology in the brain.

Figure 4. Potential models for disparate HLA haplotype effect in atypical AD versus typical amnestic AD

A. Model 1: *DR15* haplotype effect is dependent on age



B. Model 2: As a risk factor for typical amnestic AD, *DR15* is depleted in patients with atypical forms of disease.



Our study benefited from several strengths. The direct sequencing of the *HLA* region in a relatively large cohort of 745 individuals provides confidence in the results of this study and validates our previous findings from analysis of imputed genotypes. While we only validated the HLA genotypes of individuals seen at the UCSF MAC and were not able to directly sequence any samples from the ADGC cohort used in our prior study, we believe the high degree of imputation accuracy in the UCSF cohort is generalizable to the larger (n = 11,381) ADGC cohort. The cleaned dataset we used to impute our HLA genotypes was stringently quality controlled (as previously described in Chapter II) and we used very strict principal component analysis cutoffs to limit our analyses to individuals of non-Hispanic European ancestry and reduce confounding due to population stratification. Indeed, in the ADGC cohort we implemented stricter inclusion criteria than were used for our UCSF cohort, in which we used any individuals who self-reported as white. Furthermore, the concordance of sequencing to imputed genotypes we observed in the UCSF cohort is consistent to that found in other studies [2]. Thus, while we cannot directly assess the concordance of imputed HLA genotypes in the ADGC through sequencing validation, we have indirect but compelling evidence to suggest that the genotypes in this larger cohort are largely accurate (estimated at ~95% from our UCSF validation study).

Our study's use of individuals seen at the UCSF MAC provides both strengths and caveats. On one hand, we have much more detailed phenotypic information on disease presentation in individuals seen at the MAC (compared, for example, to individuals included in large national cohorts who may have had an atypical presentation but who were generally categorized as "AD" patients without any additional

phenotypic information) and the ability to verify these clinical presentations through case notes, neuroimaging and other biomarker information, and additional neuropsychological assessments. This may help explain the different effect sizes seen between the UCSF cohort (OR = 3.69; n = 309) and the ADGC cohort (OR = 1.22; n = 11,381) in our previous study, as the UCSF cohort of typical amnestic cases is likely to be relatively pure while larger national cohorts like ADGC may contain more atypical presentations under the umbrella diagnosis of 'AD'. Given our findings of the protective effect of this haplotype in atypical AD, the potential inclusion of these cases in large AD cohorts may be affecting ORs derived from these studies. However, the MAC is a clinical center known for expertise in early and atypical forms of neurodegeneration, so even amongst patients diagnosed with typical amnestic AD, the age of onset of MAC patients tends to be earlier than in community-based studies. We suspect that large national studies may have a higher frequency of atypical cases included within the umbrella AD diagnosis, whereas the present study is likely to have a higher frequency of early-onset cases included with amnestic AD diagnoses.

This study further corroborates the role of *HLA* variation as a genetic contributor to AD risk. Our novel findings of a differential effect of HLA haplotypes in typical versus atypical AD suggest the role of this region as a modifier of neurodegenerative disease may be more complex than initially thought. We plan to expand these studies through the use of additional data, including neuroimaging, gene expression studies, detailed family histories, and biomarker data from the MAC and other publicly available datasets. We will also corroborate our findings through the use of larger cohorts and the addition of other disease phenotypes. In particular, our sample sizes in this study were too small

to fully assess HLA associations with FTD. While we did find several alleles significantly associated with bvFTD or svPPA risk at the locus level, we are cognizant of the very small sample sizes for these preliminary studies and aim to expand these findings through sequencing and imputation of additional samples. As the results of this study support the use of imputed HLA genotypes in disease association studies, we plan to expand our sample size of individuals with FTD by imputing HLA genotypes from 3,526 FTD patients and 9,402 controls included in prior FTD GWAS [28] through an on-going collaboration with Dr. Raf Ferrari (UCL). This GWAS previously implicated *HLA-DRA* and *HLA-DRB5* in FTD risk through single SNP analysis. We expect that utilizing this large cohort will enable us to fine-map specific HLA haplotypes and loci underlying this association of the HLA class II region with FTD risk. We also plan to use diverse cohorts beyond white individuals of European descent to fine-map the risk associations we've previously discovered to particular alleles, as has been done for the role of HLA risk associations in multiple sclerosis [35].

In the present study, we also find continuing evidence supporting a pan-neuronal effect of particular HLA alleles in disease susceptibility. In addition to *B*07:02* and *DR15*'s effects in numerous neurodegenerative diseases (described in Chapter II), the novel finding of *C*07:02* association with greater risk of typical amnestic AD and protection from atypical AD is of particular interest given the finding that the *B*07:02~C*07:02~DRB5*01~DRB1*15:01~DQA1*01:02~DQB1*06:02* haplotype is associated with PD risk [36]. The association of *C*07:02* with this haplotype is unsurprising given the high LD (0.88) between *B*07:02* and *C*07:02*, as well as a previous study implicating *B*07:02* and *C*07:02* in AD risk particularly in individuals

without the *APOE-ε4* allele [37]. This study identified *B*07:02* and *C*07:02* homozygosity as even more strongly associated with AD risk, which we plan to explore in our cohort in the future. The *C*07:02* allele was also implicated in our analysis of bvFTD risk, which we are particularly interested in exploring in future studies given the potential implication of this allele across PD, AD and bvFTD. Locus-level analyses support the role of the *B*07:02*, *C*07:02*, *DRB1*15:01,* and *DQB1*06:02* alleles in both greater AD risk, consistent with our previous findings, and decreased atypical AD risk. The effects of the *A*02:01* allele in AD risk are conflicting [38] as previously described, and we were not able to find any studies evaluating the effect of the *A*03:01* allele in AD risk besides our own prior study [1]. While the *DPB1*04:01* allele is present in all of our significant haplotypes, we believe this locus is unlikely to underlie disease risk given that *DPB1*04:01* is present in > 60% of the European population [39], further reinforcing the importance of the *B*07:02*, *C*07:02*, *DRB1*15:01,* and *DQB1*06:02* alleles. Taken together, these findings suggest that adaptive immunity plays a critical role in maintaining neuronal health in aging. Future work will be required to elucidate the specific molecular interactions underlying this complex network of interactions in brain health and disease, and to identify potential inroads for immune-based disease-modifying therapies in neurodegeneration.

References

1. Steele NZ, Carr JS, Bonham LW, Geier EG, Damotte V, Miller ZA, et al. Fine-mapping of the human leukocyte antigen locus as a risk factor for Alzheimer disease: A case-control study. PLoS Med. 2017;14. doi:10.1371/journal.pmed.1002272

2. Karnes JH, Shaffer CM, Bastarache L, Gaudieri S, Glazer AM, Steiner HE, et al. Comparison of HLA allelic imputation programs. PLoS One. 2017;12. doi:10.1371/journal.pone.0172444

3. Pappas DJ, Lizee A, Paunic V, Beutner KR, Motyer A, Vukcevic D, et al. Significant variation between SNP-based HLA imputations in diverse populations: the last mile is the hardest. Pharmacogenomics J. 2017; 1–10. doi:10.1038/tpj.2017.7

4. Dubois B, Feldman H, Jacova C. Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria. Lancet Neurol. 2014;13: 614–29. doi:10.1016/S1474-4422(14)70090-0

5. Crutch SJ, Schott JM, Rabinovici GD, Boeve BF, Cappa SF, Dickerson BC, et al. Shining a light on posterior cortical atrophy. Alzheimer's Dement. 2013;9: 463–465. doi:10.1016/j.jalz.2012.11.004

6. Gorno-Tempini ML, Hillis AE, Weintraub S, Kertesz A, Mendez M, Cappa SF, et al. Classification of primary progressive aphasia and its variants. Neurology. 2011;76: 1006–1014.

7. Mckeith IG. Diagnosis and management of dementia with Lewy bodies: third report of the DLB Consortium. Neurology. 2006;66: 1455; author reply 1455. doi:10.1212/01.wnl.0000224698.67660.45

8. Román GC, Tatemichi TK, Erkinjuntti T, Cummings JL, Masdeu JC, Garcia JH, et al. Vascular dementia: diagnostic criteria for research studies. Report of the NINDS-AIREN International Workshop. Neurology. 1993;43: 250–260. doi:10.1212/WNL.43.2.250

9. Rascovsky K, Hodges JR, Knopman D, Mendez MF, Kramer JH, Neuhaus J, et al. Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. Brain. 2011;134: 2456–2477. Available: http://www.brain.oxfordjournals.org/cgi/doi/10.1093/brain/awr179

10. Zheng X, Shen J, Cox C, Wakefield JC, Ehm MG, Nelson MR, et al. HIBAG--HLA genotype imputation with attribute bagging. Pharmacogenomics J. 2014;14: 192–200. doi:10.1038/tpj.2013.18

11. Cereb N, Kim HR, Ryu J, Yang SY. Advances in DNA sequencing technologies for high resolution HLA typing. Human Immunology. 2015. pp. 923–927. doi:10.1016/j.humimm.2015.09.015

12. Pappas DJ, Marin W, Hollenbach JA, Mack SJ. Bridging ImmunoGenomic Data Analysis Workflow Gaps (BIGDAWG): An integrated case-control analysis pipeline. Hum Immunol. 2016;77: 283–7. doi:10.1016/j.humimm.2015.12.006

13. Maiers M, Gragert L, Klitz W. High-resolution HLA alleles and haplotypes in the United States population. Hum Immunol. 2007;68: 779–788.

doi:10.1016/j.humimm.2007.04.005

14. Dickerson BC, Wolk D a. Dysexecutive versus amnesic phenotypes of very mild Alzheimer's disease are associated with distinct clinical, genetic and cortical thinning characteristics. J Neurol Neurosurg Psychiatry. 2011;82: 45–51. doi:10.1136/jnnp.2009.199505

15. Mez J, Cosentino S, Brickman AM, Huey ED, Manly JJ, Mayeux R. Dysexecutive versus amnestic alzheimer disease subgroups: Analysis of demographic, genetic, and vascular factors. Alzheimer Dis Assoc Disord. 2013;27: 218–225. doi:10.1097/WAD.0b013e31826a94bd

16. Snowden JS, Stopford CL, Julien CL, Thompson JC, Davidson Y, Gibbons L, et al. Cognitive phenotypes in Alzheimer's disease and genetic risk. Cortex. 2007;43: 835–845. doi:10.1016/S0010-9452(08)70683-X

17. Schott JM, Crutch SJ, Carrasquillo MM, Uphill J, Shakespeare TJ, Ryan NS, et al. Genetic risk factors for the posterior cortical atrophy variant of Alzheimer's disease. Alzheimer's Dement. 2016;12: 862–871. doi:10.1016/j.jalz.2016.01.010

18. Carrasquillo MM, Khan QUA, Murray ME, Krishnan S, Aakre J, Pankratz VS, et al. Late-onset Alzheimer disease genetic variants in posterior cortical atrophy and posterior AD. Neurology. 2014;82: 1455–1462. doi:10.1212/WNL.0000000000000335

19. Carrasquillo MM, Barber I, Lincoln SJ, Murray ME, Camsari GB, Khan Q ul A, et al. Evaluating pathogenic dementia variants in posterior cortical atrophy. Neurobiol Aging. 2016;37: 38–44. doi:10.1016/j.neurobiolaging.2015.09.023

20. Henry ML, Gorno-Tempini ML. The logopenic variant of primary progressive aphasia. Curr Opin Neurol. 2010;23. Available: http://journals.lww.com/co-neurology/Fulltext/2010/12000/The_logopenic_variant_of_primary_progressive.14.aspx

21. Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet. 2013;45: 1452–8. doi:10.1038/ng.2802

22. Crane PK, Trittschuh E, Mukherjee S, Saykin AJ, Sanders RE, Larson EB, et al. Incidence of cognitively defined late-onset Alzheimer's dementia subgroups from a prospective cohort study. Alzheimer's Dement J Alzheimer's Assoc. Elsevier; 2017; doi:10.1016/j.jalz.2017.04.011

23. Schmidt H, Williamson D, Ashley-Koch A. HLA-DR15 haplotype and multiple sclerosis: a HuGE review. Am J Epidemiol. 2007;165: 1097–109. doi:10.1093/aje/kwk118

24. Lysandropoulos AP, Racapé J, Holovska V, Toungouz M. Human leucocyte antigen (HLA) class I and II typing in Belgian multiple sclerosis patients. Acta Neurol Belg. 2016; doi:10.1007/s13760-016-0716-0

25. Isobe N, Keshavan A, Gourraud P-A, Zhu AH, Datta E, Schlaeger R, et al. Association of HLA Genetic Risk Burden With Disease Phenotypes in Multiple Sclerosis. JAMA Neurol. 2016;73: 795–802. doi:10.1001/jamaneurol.2016.0980

26. Hamza TH, Zabetian CP, Tenesa A, Laederach A, Montimurro J, Yearout D, et al.

Common genetic variation in the HLA region is associated with late-onset sporadic Parkinson's disease. Nat Genet. 2010;42: 781–785. doi:10.1038/ng.642; 10.1038/ng.642

27. Wissemann WT, Hill-Burns EM, Zabetian CP, Factor SA, Patsopoulos N, Hoglund B, et al. Association of Parkinson disease with structural and regulatory variants in the HLA region. Am J Hum Genet. 2013;93: 984–93. doi:10.1016/j.ajhg.2013.10.009

28. Ferrari R, Hernandez DG, Nalls MA, Rohrer JD, Ramasamy A, Kwok JBJ, et al. Frontotemporal dementia and its subtypes: A genome-wide association study. Lancet Neurol. 2014;13: 686–699. doi:10.1016/S1474-4422(14)70065-1

29. Tafti M, Hor H, Dauvilliers Y, Lammers GJ, Overeem S, Mayer G, et al. DQB1 Locus Alone Explains Most of the Risk and Protection in Narcolepsy with Cataplexy in Europe. Sleep. 2014;37: 19–25. doi:10.5665/sleep.3300

30. Ollila HM, Fernandez-Vina M, Mignot E. HLA-DQ Allele Competition in Narcolepsy: A Comment on Tafti et al. DQB1 locus alone explains most of the risk and protection in narcolepsy with cataplexy in Europe. Sleep. 2015;38: 147–151. doi:10.5665/sleep.4342

31. Torres AR, Westover JB, Rosenspire AJ. HLA Immune Function Genes in Autism. Autism Res Treat. 2012;2012: 1–13. doi:10.1155/2012/959073

32. Al-Hakbany M, Awadallah S, AL-Ayadhi L. The Relationship of HLA Class I and II Alleles and Haplotypes with Autism: A Case Control Study. Autism Res Treat. 2014;2014: 1–6. doi:10.1155/2014/242048

33. Gough SCL, Simmonds MJ. The HLA Region and Autoimmune Disease: Associations and Mechanisms of Action. Curr Genomics. 2007;8: 453–65. doi:10.2174/138920207783591690

34. Miller ZA, Mandelli ML, Rankin KP, Henry ML, Babiak MC, Frazier DT, et al. Handedness and language learning disability differentially distribute in progressive aphasia variants. Brain. 2013;136: 3461–3473. doi:10.1093/brain/awt242

35. Oksenberg JR, Barcellos LF, Cree BAC, Baranzini SE, Bugawan TL, Khan O, et al. Mapping multiple sclerosis susceptibility to the HLA-DR locus in African Americans. Am J Hum Genet. 2004;74: 160–7. doi:10.1086/380997

36. Wissemann WT, Hill-Burns EM, Zabetian CP, Factor SA, Patsopoulos N, Hoglund B, et al. Association of parkinson disease with structural and regulatory variants in the hla region. Am J Hum Genet. 2013;93: 984–993. doi:10.1016/j.ajhg.2013.10.009

37. Lehmann DJ, Barnardo MC, Fuggle S, Quiroga I, Sutherland A, Warden DR, et al. Replication of the association of HLA-B7 with Alzheimer's disease: a role for homozygosity? J Neuroinflammation. The Oxford Project to Investigate Memory and Ageing (OPTIMA), University Department of Pharmacology & Radcliffe Infirmary, Oxford, UK. donald.lehmann@pharm.ox.ac.uk; 2006;3: 33. Available: http://www.jneuroinflammation.com/content/3/1/33

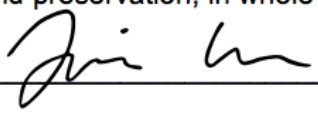
38. Cifuentes RA, Murillo-Rojas J. Alzheimer's disease and HLA-A2: Linking neurodegenerative to immune processes through an in silico approach. BioMed Research International. 2014. doi:10.1155/2014/791238

39. Sidney J, Steen A, Moore C, Ngo S, Chung J, Peters B, et al. Five HLA-DP molecules frequently expressed in the worldwide human population share a common HLA supertypic binding specificity. J Immunol. 2010;184: 2492–503. doi:10.4049/jimmunol.0903655

Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.

Author Signature 

Date 25 August 2017