PATHWAYS TO DEMENTIA:

GENETIC PREDICTORS OF COGNITIVE AND BRAIN IMAGING

ENDOPHENOTYPES IN ALZHEIMER'S DISEASE

Vijay K Ramanan

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements
for the degree
Doctor of Philosophy
in the Department of Medical and Molecular Genetics,
Indiana University

June 2013

Accepted by the Faculty of Indiana University, in partial
fulfillment of the requirements for the Degree of Doctor of Philosophy.

Andrew J. Saykin, PsyD, Chair

Tatiana M. Foroud, PhD

Doctoral Committee

Brittney-Shea Herbert, PhD

May 1, 2013

Brenna C. McDonald, PsyD

Li Shen, PhD

ACKNOWLEDGEMENTS

I thank the following people for their work, guidance, support, friendship, and example: members of my doctoral committee, researchers and staff of the Indiana University Center for Neuroimaging and the Department of Medical and Molecular Genetics, directors and students in the Indiana University Medical Scientist Training Program, and internal and external research collaborators. I also thank my family for all of these things, and their love.

ABSTRACT

Vijay K Ramanan

PATHWAYS TO DEMENTIA: GENETIC PREDICTORS OF COGNITIVE AND BRAIN IMAGING ENDOPHENOTYPES IN ALZHEIMER'S DISEASE

Alzheimer's disease (AD) is a national priority, with nearly six million Americans affected at an annual cost of $200 billion and no available cure. A better understanding of the mechanisms underlying AD is crucial to combat its high and rising incidence and burdens. Most cases of AD are thought to have a complex etiology with numerous genetic and environmental factors influencing susceptibility. Recent genome-wide association studies (GWAS) have confirmed roles for several hypothesized genes and have discovered novel loci associated with disease risk. However, most GWAS-implicated genetic variants have displayed modest individual effects on disease risk and together leave substantial heritability and pathophysiology unexplained. As a result, new paradigms focusing on biological pathways have emerged, drawing on the hypothesis that complex diseases may be influenced by collective effects of multiple variants – of a variety of effect sizes, directions, and frequencies – within key biological pathways. A variety of tools have been developed for pathway-based statistical analysis of GWAS data, but consensus approaches have not been systematically determined. We critically review strategies for genetic pathway analysis, synthesizing extant concepts and methodologies to guide application and future development. We then apply pathway-based approaches to complement GWAS of key AD-related endophenotypes, focusing on two early, hallmark features of disease, episodic memory impairment and brain deposition of amyloid-β. Using GWAS and pathway analysis, we confirmed the association of *APOE* (apolipoprotein E) and discovered additional

genetic modulators of memory functioning and amyloid-β deposition in AD, including pathways related to long-term potentiation, cell adhesion, inflammation, and NOTCH signaling. We also identified genetic associations to amyloid-β deposition that have classically been understood to mediate learning and memory, including the BCHE gene and signaling through the epidermal growth factor receptor. These findings validate the use of pathway analysis in complex diseases and illuminate novel genetic mechanisms of AD, including several pathways at the intersection of disease-related pathology and cognitive decline which represent targets for future studies. The complexity of the AD genetic architecture also suggests that biomarker and treatment strategies may require simultaneous targeting of multiple pathways to effectively combat disease onset and progression.

Andrew J. Saykin, PsyD, Chair

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AD	Alzheimer's disease
Aβ	Amyloid-beta
APOE	Apolipoprotein E
AV-45	^{18}F-florbetapir
BCHE	Butyrylcholinesterase
CNV	Copy number variant
CSF	Cerebrospinal fluid
EGFR	Epidermal growth factor receptor
EMCI	Early mild cognitive impairment
eQTL	Expression quantitative trait locus
FDR	False discovery rate
GO	Gene Ontology
GWAS	Genome-wide association study
HC	Healthy control
LD	Linkage disequilibrium
LMCI	Late mild cognitive impairment
MAF	Minor allele frequency
MAPK	Mitogen-activated protein kinase
MCI	Mild cognitive impairment
PET	Positron emission tomography
Q-Q	Quantile-quantile
SNP	Single nucleotide polymorphism

I. Introduction

A. Alzheimer's disease (AD): clinical and pathologic features

Alzheimer's disease (AD) is a progressive and debilitating neurodegenerative disorder that causes substantial loss of memory and other cognitive and social functions, eventually leading to dementia characterized by severe impairment in activities of daily living [1]. Currently, nearly six million Americans are affected with AD at an annual cost of $200 billion [2-4]. Although some drug therapies are available on a symptomatic basis, there is currently no cure for the underlying disease [5]. With its high and rising incidence and burdens in an increasingly aging population, the development of treatments to slow, halt, or reverse AD onset and progression is a national priority [4].

Pathologically, the hallmark abnormalities of AD are the accumulation of abnormal and/or misfolded amyloid-β (Aβ) peptides in the extracellular spaces of the brain and the presence of intracellular neurofibrillary tangles composed of tau protein and degenerating cellular structures [6]. These pathologic features are thought to interfere with synaptic connections between neurons as well as disrupt neuronal metabolism and energetics, leading to disease-characteristic structural and functional changes in the brain [3].

The most common early symptom of AD is impairment in episodic memory, involving the encoding of new experiences and subsequent conscious recollection of past experiences [7]. Declines in language, executive function, judgment, orientation, mental status, and emotional and social health are also frequent findings with significant implications for patients and their

caregivers [7]. These progressive and debilitating symptoms have stimulated investigation into strategies for early detection of individuals exhibiting prodromal stages of the disease characterized by milder impairment in cognitive performance and self- and informant-reports of cognitive decline [8-10]. As a result, numerous disease-related biomarkers have been proposed for potential clinical application – including measures of cognitive performance, structural and functional brain imaging, and blood and cerebrospinal fluid (CSF) analytes, among others [3] – which can also serve as quantitative disease-related endophenotypes that provide enhanced power and biological interpretability for studies of the genetic etiology underlying AD [11, 12].

B. Genetic architecture of AD

Prior to the advent of genome-wide association studies (GWAS), the discovery of mutations in *APP* (amyloid precursor protein) and *PSEN1* and *PSEN2* (presenilin 1 and 2) causing rare, early onset, familial forms of AD dominated much of the genetic research and biological theory about the disease [13]. However, most (> 90%) cases of AD do not appear to be caused by simple Mendelian inheritance of causative mutations and tend to have later onset (after 60-65 years of age). Instead, these cases are thought to have a complex etiology with multiple genetic and environmental factors contributing to susceptibility.

The strongest known genetic risk factor for AD is the *APOE* (apoliprotein E) ε4 allele. Individuals with two copies of *APOE* ε4 have up to 30 times the risk of developing AD compared to individuals without this allele [14]. However, *APOE* ε4 is neither necessary nor sufficient for development of AD or its characteristic pathology, suggesting that other genes influencing disease status remain to be discovered. Unbiased GWAS, which test up to several million

common genetic variants (population minor allele frequency > 1%) for association to disease or disease-related traits, have uncovered additional susceptibility loci which have suggested previously unpredicted disease mechanisms. These have included genes related to inflammation and immune activation, protein degradation, lipid metabolism, endocytosis, and other candidate disease mechanisms [15-26]. Unfortunately, many AD GWAS-implicated single nucleotide polymorphism (SNP) variants have displayed modest individual effect sizes and associations for some loci have not been replicated. In addition, although based on twin studies up to 60-80% of AD risk is estimated to derive from genetic factors [27], known genes including the uniquely large effect of *APOE* account for just half of this genetic variance [28].

These limitations of existing GWAS findings in AD have spurred significant interest in the development of alternative perspectives and analytical strategies to better understand the genetic architecture underlying the disease [29, 30]. In particular, biological pathways and networks have become focal points for harnessing GWAS data in complex diseases [31, 32]. Numerous studies have demonstrated that genes functioning in the same pathway can collectively influence susceptibility to neurodegenerative diseases and traits, even when constituent SNPs do not individually exhibit the robust association required for genome-wide significance [33-40]. Pathways occupied by top GWAS "hits" can also highlight additional genes with more modest effects on disease risk but which may provide better targets for biomarker and drug development [41, 42]. Further, GWAS-implicated pathways and networks provide mechanistic hypotheses which can guide confirmatory testing in independent human study data sets, cell lines, and animal models. The ability to prioritize pathways of interest may be particularly important for approaches with high computational demand. These include whole genome sequencing (WGS) studies, which offer enhanced power to detect rare SNPs and copy

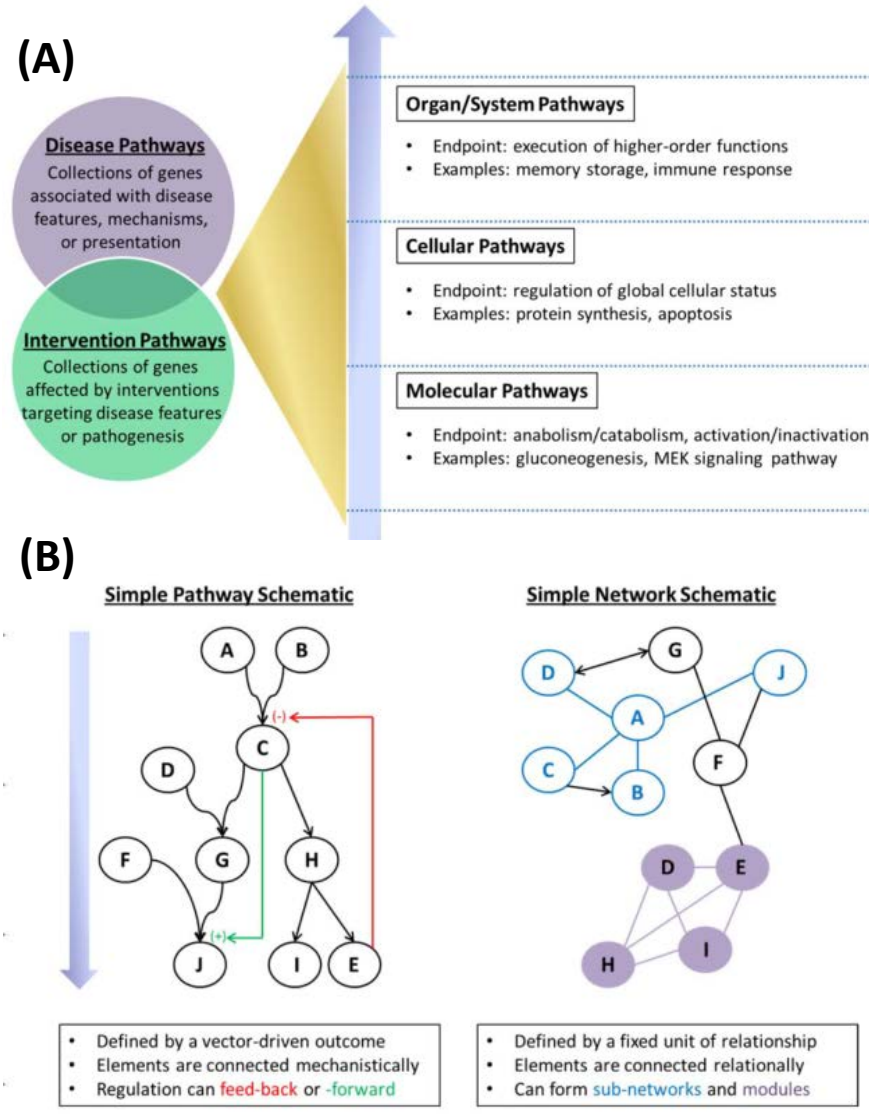
number and other structural variants [43], studies of disease endophenotypes such as brain imaging [11, 44] or cerebrospinal fluid (CSF) biomarkers [45, 46], and studies of molecular interactions and epistasis [47-49], among other approaches. Despite this potential for pathway-focused study designs to extend and validate standard GWAS approaches to AD, fundamental concepts and factors related to pathway-based study design are still new and have been evolving.

C. Fundamental concepts about biological pathways and networks

While unstated notions predate it, the first explicit description of a pathway as the events by which intermediates are processed in a defined sequence was provided in 1973 [50]. Recently, broader notions of pathways as collections of biologically-related genes [51] have attempted to fit evolving scientific theories and analyses. A more systematic conceptualization of biological pathways (Figure 1A) can be achieved on the premise that pathways are vector-driven toward an essential goal (i.e., their constituents as a whole are directed to a common, specific endpoint). Viewed this way, molecular pathways have an essential goal of basic biochemical action on molecules or compounds. Meanwhile, cellular pathways regulate global cellular status, and organ/system pathways execute broader physiological functions. The constituents of these pathways are typically connected through known or proposed mechanisms. Of note, the particular constituents of a pathway may be context-dependent – specifically, in relation to the biological outcome an investigator wishes to study.

In addition, two other types of pathways are important in the study of genetically-complex diseases (Figure 1B). Disease pathways have an essential goal of the pathogenesis of a disease

Figure 1. A primer on biological pathways and networks. (A) The major types of biological pathways are shown along with a representation of their relationships among each other. Each type of pathway is defined by its essential goal. (B) Pathways can include directional regulation (shown in red and green), branching, and mechanistic connections leading to an essential outcome. Network elements are connected through shared relationships and are not vector-driven from a starting point to an essential outcome. Networks can be divided into subnetworks (shown in blue) exhibiting all elements connected to a central node ("A" in this example) or into modules (shown in purple) that exhibit a high density of connections.



and its features. For example, the AD pathway plausibly includes components from the organ/system pathway of memory, which itself has cellular and molecular underpinnings. Meanwhile, intervention pathways are defined within the setting of a therapy that targets disease features or pathogenesis, as in a pathway-based study of cisplatin sensitivity in ovarian cancer [52]. Importantly, disease and intervention pathways may include constituents with documented associations to a phenotype, but whose precise mechanistic roles are not yet known.

Networks can also collect genes and other biological elements for quantitative and visual assessment of relationships [53]. Unlike pathways, biological networks are not vector-driven toward an essential outcome (Figure 1B). Instead, networks are characterized by nodes that are connected by edges representing defined relationships. In a particular network, nodes may represent nearly any biological element, including genes, gene products, non-gene DNA sequences, pathways, diseases, therapies, or combinations thereof. Common examples of network relationships include binding in protein interaction networks and regulation by common factors in gene interaction networks. Finally, statistical networks display relationships, such as correlation, that are inferred from computational analyses [54]. A central outstanding question involves understanding the degree of connection between statistically-inferred networks and biological networks [55]. Software platforms for network analysis include IPA (Ingenuity Systems) and Cytoscape [56]; two recent reviews discuss these and other network-based tools in detail [57, 58].

Despite their popularity and potential, strategies for pathway-based genetic studies have largely progressed in the absence of guidelines. This has led to ambiguity regarding optimal methods

and high variability in results, creating challenges for biological interpretation and barriers to further application. In particular, at the time of commencing this work, only one published report described a pathway analysis of GWAS data on AD susceptibility [59], and no genome-wide pathway analyses had been reported for quantitative AD-related endophenotypes. These findings, together with the encouraging results from pathway analyses in other biological realms including breast cancer [60], Crohn's disease [32], type 2 diabetes [61], and multiple sclerosis [38, 39, 62], suggested that pathway analysis might serve as a powerful complementary approach for confirming existing hypotheses and discovering novel mechanisms contributing to AD and its related endophenotypes.

D. Statement of purpose

With the rising incidence and burdens of AD, a better understanding of its underlying genetic mechanisms is crucial for the development of effective diagnostic and therapeutic strategies. Although GWAS of AD and AD-related endophenotypes have been fruitful, the focus to-date on individual susceptibility variants has raised several key limitations and has left substantial heritability and pathophysiology unexplained by extant findings. Pathway-based analysis represents a promising approach for extending the utility of GWAS in AD to identify novel influences on disease, but at the time of commencing this work, the lack of a systematic framework for conducting pathway-focused studies had limited the application of these analytical techniques.

Accordingly, the overall goals of this work were as follows:

1. Critically review strategies for genetic pathway analysis, synthesizing extant concepts and methodologies to guide application and future development.
2. Perform pathway analysis to complement GWAS of endophenotypes representing the early, hallmark AD features of episodic memory impairment and cerebral Aβ deposition.

We hypothesized that pathway analysis would confirm candidate mechanisms nominated by previous GWAS in AD as well as highlight novel biological mechanisms that are otherwise concealed from standard GWAS approaches. This discovery of novel biological pathways underlying AD would be crucial for the development of diagnostic and therapeutic approaches to slow, halt, and reverse disease onset and progression. In addition, we expected that this work would provide a systematic structure for genetic pathway analysis that would enhance its application and help validate its promise for future studies of complex diseases and traits.

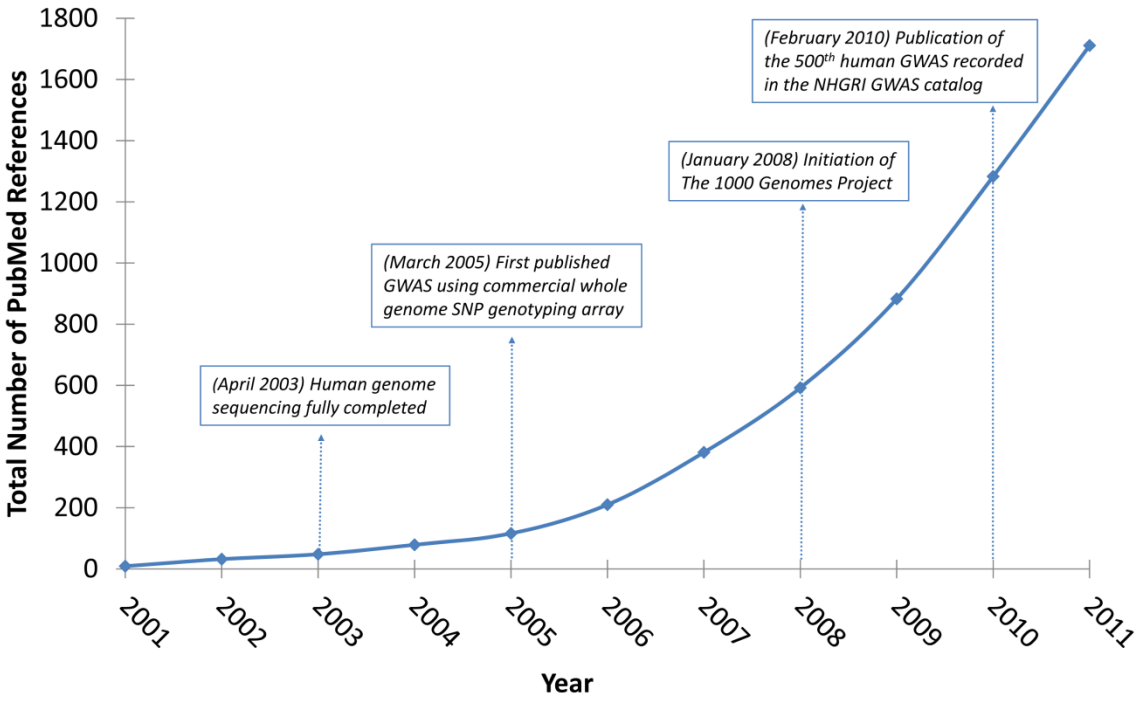
II. Pathway analysis of genomic data: concepts, methods, and prospects for future development

A. Introduction

Since 2005, over 1000 human GWAS publications have described genetic associations to a wide range of diseases and traits [63]. However, extending GWAS findings to mechanistic hypotheses about development and disease has been a major ongoing challenge. In particular, the focus on single loci has been confounded by two insights: 1) most GWAS-implicated common alleles and differentially-expressed genes on expression arrays have exhibited modest effect sizes; and 2) genes function within biological pathways and interact within biological networks [49]. As such, genome-wide data sets are increasingly viewed as foundations for discovering pathways and networks relevant to phenotypes [41]. This trend is vital, given that pathway mechanisms are natural sources for developing strategies to diagnose, treat, and prevent complex diseases. In this context, it is not surprising that pathway-based analyses have exploded in use during the last 3-5 years (Figure 2).

In pathway analysis, gene sets corresponding to biological pathways are tested for significant omnibus relationships with a phenotype. Primary data for pathway analysis is commonly sourced from genotyping or gene expression arrays, though in theory any data elements that could be mapped to genes or gene-related products could be used. Importantly, analyzing genomic data through functionally-derived gene sets can reveal larger effects that are otherwise concealed from gene- or SNP-based analysis. For example, high-profile studies in breast cancer [60], Crohn's disease [32], and type 2 diabetes [61] demonstrate that functionally-related

Figure 2. PubMed citations for "pathway analysis": 2001-2011. The use of pathway analysis has grown exponentially in the last 3-5 years. This explosion in use has followed major developments (shown in boxes) in characterizing the human genome and in performing genome-wide studies of complex diseases and traits. Data points represent the total number of references displayed through a PubMed search for "pathway analysis", using date limits of January 1, 2001 and December 31 of the calendar year denoted on the x-axis.



genes can collectively influence disease susceptibility, even if individual loci do not exhibit genome-wide significant association in a particular data set. As such, pathway analysis represents a potentially powerful and biologically-oriented bridge between genotypes and phenotypes.

Despite their popularity and potential, strategies for pathway-based studies have progressed in the absence of guidelines, leading to ambiguity regarding optimal methods, high variability in results, and barriers to further application. With surging interest in pathway analysis and the emergence of next-generation sequencing data which will inevitably broaden its application, this is an ideal moment for a critical synthesis of current approaches to guide application and future development. Here, we review extant strategies to detect pathway-phenotype association, highlight methodological considerations and challenges, and describe how pathways and networks are ideal vehicles for leveraging multi-omics data for discovery.

B. Selecting an overall study design

Broadly, there are two approaches to pathway-based genomic studies. Candidate pathway analysis is hypothesis-driven: pathways are preselected based on prior knowledge and insight. While the number of candidate pathways may vary with study goals (e.g., different effects may be seen within a large, complex pathway compared to numerous, smaller pathways), this approach is marked by its use of a biologically-targeted subset of genomic data. In contrast, genome-wide pathway analysis interrogates a complete genomic data set through pathways representing an extensive range of biology. Notably, the line between "targeted" and "extensive" biological coverage is not precisely drawn. While methods limited to genome-wide

pathway analysis have been used on data sets with only 1000 genes (~5% of the total number of human genes) [64], the optimal point of delineation between these two approaches warrants further examination.

There are several advantages to the candidate pathway approach. Focusing the scope of analysis can enable otherwise intensive procedures like genotype imputation and manual pathway curation. By maximizing annotation coverage and quality, these procedures can bridge differences in genotyping platforms across cohorts for replication or meta-analysis. Unfortunately, targeted biological coverage may fail to detect unexpected relationships, such as the association between inflammatory pathways and age-related macular degeneration [65]. Further, poor annotation of one pathway can be particularly limiting when only a few pathways are assessed. These traits make candidate pathway analysis most appropriate where computational resources are a consideration and where specific pathways are of *a priori* interest.

In contrast, genome-wide pathway analysis maximally utilizes the available genomic data. As a result, this approach can more readily detect unexpected relationships, including those across diseases operating in different body systems [66]. However, genome-wide pathway analysis is computationally intensive, requiring more stringent corrections for multiple comparisons and making covariance analyses, imputation, and meta-analyses more challenging. While strategies to reduce the dimensionality of genome-wide data for pathway analysis are in active development [67, 68], they will need to be evaluated further ahead of widespread use. Genome-wide pathway analyses also benefit from systematic follow-up to deal with the often

high overlap of genes across multiple pathways and to evaluate results in view of prior knowledge.

C. Obtaining input genomic and pathway annotation data

Pathway analyses can utilize raw genotype data for individual subjects [61, 69, 70] or a list of *p*-values relating genes or SNPs to a phenotype [71-73]. Most pathway-based tools for raw genotypes do not effectively include covariates but can naturally correct for linkage disequilibrium (LD) through permutation. In contrast, *p*-value distributions are readily accessible via other researchers and can be generated through covariance analyses, but require corrections for LD based on reference population data. Investigators should consider their resources and study goals when selecting the most appropriate genomic data source.

In parallel, a pathway analysis is only as good as the functional information underlying its pathway definitions. Prominent pathway annotation databases exhibit diverse features (Table 1; also see the online resource Pathguide [74]). The ideal choice of database depends on several variables and their impact on study goals. For example, publically-available freeware databases are commonly used due to their ease of access, transparency of features, and visibility in publications. Commercial access databases may require a significant investment; however, they are typically linked to user-friendly statistical analysis software and often include high-quality pathway graphics which can be exported to manuscripts. Investigators should weigh the relative importance of these factors during selection.

Table 1. Prominent pathway annotation databases.

Name	Curation[a]	Major Features	URL
Biocarta	M	Driven by user input; some expert review	www.biocarta.com
DAVID	M/E	Augments and integrates annotations from other databases	david.abcc.ncifcrf.gov
Gene Ontology	M/E	Largest database; hierarchical structure; can filter data by evidence codes	www.geneontology.org
Ingenuity	M/E	Large collection of canonical pathways; high-quality pathway maps	www.ingenuity.com
Kyoto Encyclopedia of Genes and Genomes	M	Reference pathways (mosaics from several organisms) and organism-specific annotations; pathway maps link to closely-related genes	www.genome.jp/kegg
MetaCore	M	Extensive disease pathways; can edit pathway maps for publication	www.genego.com
MetaCyc	M	Metabolic pathways; can visualize connections among pathways	metacyc.org
Molecular Signatures Database	M/E	Can download pathways from several other databases as a collection for input to analytical software; novel motif gene sets	www.broadinstitute.org/gsea/msigdb
PANTHER	M	Can predict protein functions from sequence and evolutionary data	www.pantherdb.org

Pathway Interaction Database	M/E	Broad range of cellular pathways with special focus on cancer signaling; can generate interaction maps from a list of genes	pid.nci.nih.gov
Reactome	M	Pathways are extensively cross-referenced to PubMed, HapMap, and other resources; can overlay expression or other data onto pathway maps	www.reactome.org
ResNet Series	M/E	Regular updates through web server; optional user editing or text scanning of user documents; links to reference articles	www.ariadnegenomics.com

[a]Abbreviations: M = manual, M/E = manual and electronic

Pathway curation methods can also impact analyses. Most databases rely on expert review for pathway curation; however, users of these databases should be aware of their update intervals and criteria used as evidence for inclusion in pathways. Alternatively, electronic curation employs text-searching algorithms to infer functional relationships. While these inferred annotations can be useful for hypothesis generation, their accuracy is unreliable [75], making them unsuited to many pathway analyses. Finally, targeted manual curation can be particularly appropriate when an investigator has expertise in a biological realm that is poorly annotated in databases. While potentially time-consuming, manual curation can synthesize recent results with established relationships to produce novel candidate pathways [76, 77] or gene sets representing positive controls for pathway analysis [78].

Lastly, the biological coverage of pathway annotations should be considered. Across databases, similarly-named pathways can exhibit vast differences in constitution while differently-named pathways can exhibit significant overlap. As a result, investigators should attempt to match study goals with database coverage. For example, specialized, high-granularity databases are most useful for candidate studies of intricate signaling pathways, while canonical pathway collections (representing well-established pathways) provide a broad biological scope well-suited for screening-oriented studies.

This collective diversity of features is a major factor in explaining why different databases can yield divergent results from the same input data [79]. As such, an early discussion of pathway analysis recommended the use of multiple databases for each analysis [30]. This approach can balance the relative characteristics of each database used and can yield a measure of validation when different databases yield similar results. However, for some genome-wide pathway

analyses, the use of a single, comprehensive database may be an optimal fit for a given study design and may facilitate straightforward interpretation of findings. In either case, a systematic qualitative review of the results is crucial to identify robust relationships and extract overarching biological themes. In addition, follow-up analyses can reveal broader findings that drive association signals across multiple smaller pathways, as with one study that analyzed pathway sets obtained through hierarchical clustering and identified an association between the canonical RAS/RAF/MAPK signaling pathway and breast cancer [60].

D. Preparing data for association testing

Systematic processing of input genomic data and pathway annotation data are vital for pathway analyses. While some relevant methods are actively evolving, optimized approaches to major issues can minimize variation in results and interpretation.

Pathway size

Most pathway analyses place constraints on pathway size: small pathways can exhibit false positive associations due to large single-gene or single-SNP effects [51], whereas large pathways are more likely to show association by chance alone [79]. The most common minimum threshold for pathway size appears to be 10 genes [60, 61, 70, 80]. It is important for analysts to note that this threshold may exclude highly-specific and potentially-informative functional sets, including those involving protein complexes and DNA sequence motifs. Frequently-used maximum thresholds for pathway size include 100 genes [60] and 200 genes [61, 80]. Notably, in the latter two studies, upper limits of 300 genes [61] and 400 genes [80] did not alter the

results. However, larger pathways are relatively rare and often derive their size from being more general in scope; thus, their exclusion may not significantly affect analyses or downstream biological interpretation. Overall, investigators should consider their study goals when applying such thresholds and should evaluate results in that context. While future efforts might develop size-dependent statistical corrections, at present the reporting of pathway size and related summary statistics (e.g., [81]) alongside association data can aid interpretation.

Pathway overlap

Genes and their products typically act in multiple pathways [49], and each role is potentially important to a disease or treatment mechanism. As a result, analyses can expect to have some degree of pathway overlap. However, high pathway overlap can obscure the true source of an association signal. While this problem can exist with any pathway analysis, Gene Ontology (GO) annotations are particularly susceptible due to the database's large, hierarchical structure [82]. Some studies have restricted analysis of GO terms to certain levels in the hierarchy [70, 83], while a new Bayesian method incorporates the structure of the hierarchy as prior information into its pathway association metric [84]. However, users of these approaches should be aware that the information content at particular GO levels is unpredictable [85]. Pathway overlap can also be addressed during post-analysis to prioritize related pathways for further exploration. Extant strategies include hierarchical clustering in a study of breast cancer [60], overlap-based network creation in the visualization tool Enrichment Map [86], and the listing of overlapping pathways alongside results in the analytical software PARIS [87].

Assigning data elements to genes

Genomic data has historically been integrated into pathways by mapping assayed elements to genes. For SNP-based genotyping arrays, this is not straightforward because many array SNPs are not located in known coding or regulatory regions. One solution discarded all SNPs that were not mapped to a single gene through a reference genome build, but this resulted in a loss of more than 25% of assayed SNPs [88]. Alternatively, each unmapped SNP can be assigned to its nearest gene [39]. However, evolving theories suggest that sequences may not be associated to genes based on closest proximity, and may not even be solely associated to one gene [89, 90]. Hence, many studies assign unmapped SNPs to all genes within a distance window, ranging from 10 kb to 500 kb [70, 80, 81, 91]. Studies taking this approach should beware that some SNPs may not be functionally related to their assigned gene(s). In addition, SNPs that map to multiple genes in the same pathway can yield spurious pathway association. This issue is particularly important for genes (such as the MHC/HLA genes) that cluster in the genome and belong to the same pathway, because variants in those genomic regions can potentially map to all genes in the pathway. Finally, given the importance of SNP-to-gene mapping for pathway analyses, investigators should be aware that imputation can increase genomic coverage by probabilistically-predicting SNP genotypes that are not directly available in a particular data set. Imputation can be particularly useful for bridging differences in genotyping platforms across cohorts for replication and meta-analysis, and can also enable investigation of rare alleles and copy number variants (CNVs) that are less-represented on standard platforms [92].

Calculating gene significance and accounting for LD

Most pathway analysis tools utilize one association signal per gene. While expression arrays yield a single *p*-value for each gene, SNP arrays include multiple signals per gene, some of which are correlated. As such, some studies use the minimum SNP-level *p*-value within a gene as the operative signal [39, 60, 80, 88]; however, this approach will not detect additive effects among SNPs with moderate individual association. For methods that combine SNP-level signals, such as those based on the truncated product method [71], LD must be accounted for to prevent highly-correlated SNPs from biasing gene-level significance. Strategies to accomplish this include discarding SNPs that depart from LD at a preset threshold [80, 81, 93] as well as adapting principal component analysis to extract the most independent signals within a gene [67, 68, 81]; unfortunately, these methods can eliminate substantial information. Alternatively, the SNP ratio test [37] and the "set-based analysis" in PLINK [94] use phenotype permutation to naturally correct for biases introduced by LD and gene size; however, these tools require raw genotype data and are computationally demanding, making them better suited for studies of candidate pathways with relatively few genes. Notably, recently-developed methods that accept *p*-values as input and account for LD through simulations [95, 96] or genotype permutation [87] are computationally efficient and may represent new paradigms as their power is honed and evaluated.

E. Analytical methods to detect pathway-phenotype relationships

Following data processing, analytical methods can be applied to test for significant pathway-phenotype relationships. Prominent examples of pathway-based analytical tools and their

salient features are provided in Table 2. Notably, one class of tools employs text-mining of published abstracts to identify potential pathway-phenotype relationships. These tools query a list which may include SNPs meeting a *p*-value threshold, genes from candidate pathways, or pathways themselves, among other possibilities. Text-mining approaches have efficiently identified potential interactions among genes associated with neurodegenerative brain changes [77] and have equally been applied to generate a candidate pathway based on regulation or interaction with *BRCA2* (breast cancer 2, early onset) [97].

By contrast, pathway enrichment tools assess for a statistically-significant distribution of association within a pathway. Competitive enrichment methods compare the collective association within a pathway to the collective signal among genes not in the pathway [98]. As a result, competitive methods are not suitable for candidate pathway analyses that do not have an appropriate complement of data from outside of the candidate pathways. Meanwhile, self-contained enrichment methods test the signal within a pathway against simulated data sets which are expected to have no significant phenotype association [98, 99]. Self-contained methods can be challenging to use in a screening-oriented genome-wide pathway analysis due to the computational demand of generating simulated data sets. In addition, self-contained approaches are particularly susceptible to false positives through genomic inflation (systematic increases in GWAS test statistics due to population stratification or other confounding factors [100, 101]), as each pathway is evaluated independently from any other data on the source assay. While one study [102] normalized all association statistics to a genomic inflation factor (λ) calculated by PLINK, best practices in this area have not yet been settled. Competitive tests are more robust in controlling genomic inflation, but they can also relinquish power in data sets

Table 2. Examples of publically-available pathway-based analytical tools.

Name	Type[a]	Input Data	Analytical Method	Corrections Included	Ref
Chilibot	TM	Word List	Searches PubMed abstracts for relationships among word list; can distinguish biological concepts (e.g., activation, inhibition)	N/A	[103]
GenGen	C	Raw genotype data	Uses best *p*-value as gene-wide score and calculates rank-based Kolmogorov-Smirnov-like pathway statistic with permutation	LD, pathway size, gene size, FDR	[104]
GeSBAP	C	Gene or SNP *p*-values	Uses best *p*-value as gene-wide score and performs rank-based Fisher's exact test to detect pathway enrichment	FDR	[105]
GRAIL	TM	SNPs or genomic regions	For multiple disease-associated regions, identifies functionally-related genes that likely highlight causal pathways	Number of genes per region	[106]
GRASS	SC	Raw genotype data	Uses principal component analysis to select representative eigenSNPs for each gene for pathway-based ridge regression	LD, gene size, FDR	[107]

GSA-SNP	C	SNP *p*-values	Uses -log(*k*th best *p*-value) as gene-wide score and calculates a z-score, iGSEA, or MAXMEAN statistic for the pathway	Pathway size, FDR	[108]
GSEA-P	C	Gene *p*-values	Calculates rank-based Kolmogorov-Smirnov-like pathway statistic with phenotype permutation	LD, pathway size, FDR	[109]
GSEA-SNP	SC	Raw genotype data	Uses all SNPs for a pathway MAX-test (maximum of Cochran-Armitage trend tests under 3 genetic models) with permutation	LD, pathway size, gene size, FDR	[110]
MAGENTA	C	SNP *p*-values	Modified approach based on GSEA-SNP for meta-analytic data	LD, gene size, FDR	[111]
PARIS	SC	SNP *p*-values	Identifies the significant genomic features within a pathway and performs genomic permutation to assess pathway significance	LD, pathway size, gene size, FDR	[87]
PLINK set test	SC	Raw genotype data	For SNPs passing a *p*-value threshold, calculates the average test statistic for the independent SNPs within a pathway	LD, pathway size, gene size, FDR	[94]

SNP Ratio Test	SC	Raw genotype data	Calculates the ratio of significant SNPs to all SNPs in a pathway and uses phenotype permutation to calculate empirical *p*-value	LD, pathway size, gene size, FDR	[37]

[a]Abbreviations: TM = text-mining, C = competitive enrichment, SC = self-contained enrichment

with diffuse association signal [98]. As such, the optimal method depends on study goals, data set properties, and computational resources.

Among extant competitive enrichment methods, three analytical frameworks predominate. In the first of these, threshold-based approaches, hypergeometric, chi-square, or Fisher's exact test statistics are used to identify pathways that are overrepresented among the "significant" markers under study. Notably, the threshold for "significance" is arbitrary and can affect results [33]; observed SNP-level thresholds have ranged from $p < 0.05$ [91] to $p < 5 \times 10^{-8}$ [39]. In contrast, rank-based approaches order all of the markers being studied by their significance and then test for pathways which have lower rankings than the overall distribution. While the rank-based tools GenGen [104] and GSEA-SNP [110] use a Kolmogorov-Smirnov-like running sum that gives greater weight to more significant markers, others rely on MAXMEAN-related statistics as potentially powerful and efficient alternatives [108, 112, 113]. Compared with threshold-based methods, rank-based approaches more naturally account for differences in significance among markers [51] but may also be heavily influenced by a few highly-significant markers [114]. Finally, z-score methods infer enrichment based on deviation from a normal distribution that accounts for the size of each pathway [108, 115]; while these methods are sensitive and fast, their error rates have not been well-characterized. Self-contained enrichment methods employ even more diverse statistical methods to combine the *p*-values within a pathway into an aggregated measure (Table 2). However, in the absence of large-scale power comparisons among related methods across several well-characterized data sets, the choice of a particular enrichment tool may be less important than understanding the relative strengths and limitations of these broader categories.

An alternative to enrichment methods are network-based approaches, which examine sets defined by other biological characteristics for meaningful pathways contained therein. For example, one study used hierarchical clustering to form networks of co-expressed genes across multiple inflammatory diseases; subsequent analysis of these networks suggested a role for interferon-inducible signaling in tuberculosis [116]. Gene networks can also be defined through protein interactions, as in a study that associated genetic variants in glutamate pathways to brain glutamate concentration in multiple sclerosis [117]. Importantly, recent studies are combining enrichment and network-based methods to point to broader findings. For example, network analysis of enriched pathways revealed major roles for antigen presentation and interferon signaling in rheumatoid arthritis [118].

Finally, developing strategies are targeting specific pathway-based challenges. For example, machine learning approaches [68, 119] attempt to identify the most informative subsets of genes within pathways for association. Networks have been effective in studies of rare variants, as with the identification of a synaptogenesis gene network affected by rare CNVs in autism [120]. Pathway-based methods for studying rare variants using genomic-region-based mapping and self-contained tests are also evolving [121, 122]. Indeed, the appeal of pathways and networks will continue to expand as their associated tools progress to analyze a variety of data through user-friendly platforms.

F. Post-analysis considerations

Following pathway analysis, appropriate data reporting and interpretation are imperative. Currently, bias introduced by gene size is less commonly addressed than bias from pathway size.

In particular, large genes containing many SNPs are more likely to contain significant SNPs by chance alone [123]; for analyses, this can favor pathways containing large genes. Analytical tools that employ permutations naturally control for gene size by comparing the actual association data to the distribution of association statistics generated from randomly permuted data sets expected to reflect chance-based confounding effects. Other approaches [94, 95] allow users to restrict analysis to a subset of the most significant SNPs in each gene: for large genes, this may eliminate some spuriously-associated SNPs and thus limit their impact on the pathway analysis. At minimum, studies should discuss potential impacts of gene and pathway size on their results. Other sources of bias that should be addressed include the capacity for strongly-associated markers to drive pathway association and the possible effects of SNPs being assigned to multiple genes.

Correction for multiple comparisons must also be applied to pathway *p*-values to control for false positives. As in other areas of statistical genomics, optimizing methods for correction is a work in progress. Bonferroni-related methods seem too conservative for pathway analyses because they do not allow for dependence across pathways. False discovery rate (FDR) approaches [124] are frequently-applied in pathway analyses [33, 61, 81], while newer FDR-based [125] and bootstrapping [93] methods that assess the uncertainty of statistical estimates through permutation can better account for pathway overlap but require large computational capacity.

Fundamentally, these approaches to bias are best complemented by replication of pathway analysis findings in independent data sets. Strategies for pathway analyses can flexibly adapt to differences across data sets, and while these differences might impact SNP- or gene-level

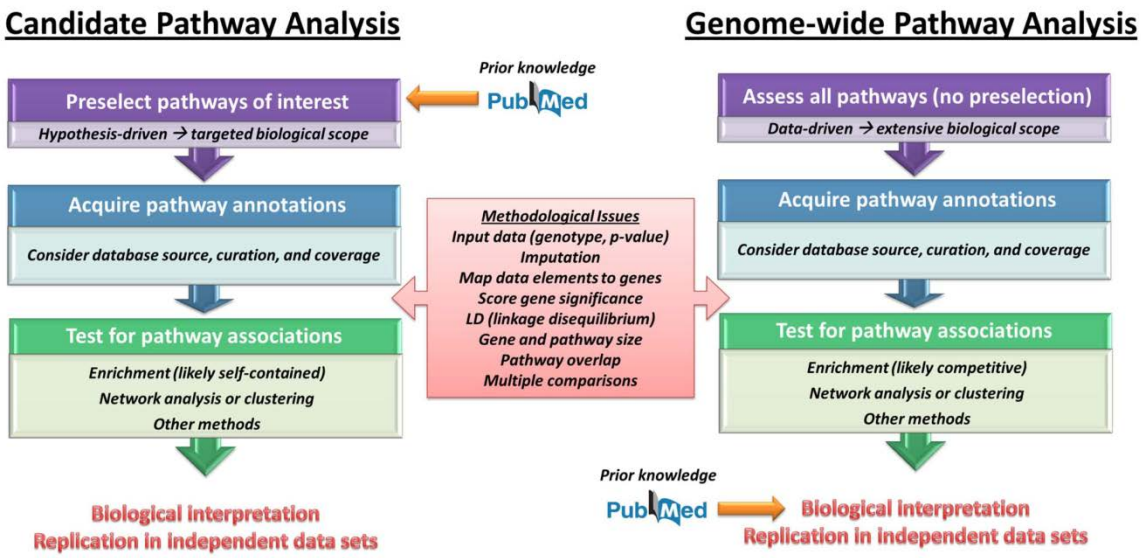
statistics [126], legitimately-associated pathways would be expected to exhibit significance or a strongly-trending signal across multiple studies. In this effort, a systematic framework illustrating key choices in pathway analyses (Figure 3) will limit major contributors of variance across studies and will guide investigators in selecting approaches that fit their study goals.

G. Future developments in genomic data analysis

Development of methods and tools related to pathway analysis is ongoing and dynamic. In particular, because pathways are of broad interest, targeted adaptations to their associated databases would expand their utility for investigators from a variety of backgrounds. These adaptations might include simpler search and download mechanisms, consistency in pathway names and classifications, and methods for describing pathway overlap. In addition, a universal format for annotation files might encourage interoperability among analytical tools, allowing investigators flexibility to precisely match their databases and statistical methods of choice.

Two recent trends among databases are also promising. Specialized disease databases, such as AlzGene [127], PDGene [128], and the UCSC Cancer Genomics Browser [129], can aggregate salient information from diverse studies on a particular disease. These targeted resources are particularly up-to-date and can facilitate collaboration within highly-investigated diseases. Functional annotation of genes is also becoming prominent. These annotations draw on experimental data that indicates function, location of action, or physiological region of association [130], and can allow investigators to develop candidate pathways related to localized anatomical or physiological derangements. Extensions of this concept across disciplines will likely be a prime area of advancement.

Figure 3. A methodological guide to pathway analysis. Broadly, there are two approaches to pathway analysis. In candidate pathway analysis, prior knowledge is used to select pathways hypothesized to have a relationship with a phenotype. In contrast, genome-wide pathway analysis is designed to uncover significant pathway-phenotype relationships within a large data set; insight and prior knowledge are then used to interpret the findings. In both approaches, care must be taken in acquiring pathway annotations and in selecting an appropriate analytical test for association. In addition, other methodological issues (red box) guide the choice of approach and impact strategies for confounding factors. Finally, replication of pathway analysis findings in independent data sets is imperative in validating results to extend their impact.



In future pathway analysis platforms, computational efficiency will be highly-valued due to the impressive granularity of next-generation sequencing data. In addition, investigators may wish to use different genomic data sets, pathway annotation databases, and analytical parameters depending on study resources and goals; as such, tools that are flexible to various study approaches will maximize their impact. Finally, given that genes constitute only 1-2% of the human genome, strategies to leverage both genic and non-genic data for pathway analysis may provide increased power to detect meaningful functional sets.

Meanwhile, complementary methods can extend the biological reach of pathway-based results. For example, it is not yet understood whether gene interactions are more likely within a given pathway or across different pathways in a network. A comparative study of epistasis in pathways and networks, perhaps utilizing novel techniques for its detection within population data [47, 54, 131, 132], could inform future strategies in this area. A related area of development involves using known protein interactions to generate subnetworks from enriched pathways; these subnetworks can highlight novel candidate genes [133] or regulatory relationships [134] from significant pathways.

Nevertheless, the ongoing development of pathway-based tools would benefit from further empirical evaluation of current approaches. For example, a creative meta-analysis might examine how various association metrics affect the likelihood of replication of findings. In addition, testing association methods against well-calibrated positive and negative control data sets might illuminate their relative capabilities. Notably, one study employed multiple pathway analysis algorithms using an extensively-explored Crohn's disease data set [135]; however, the algorithms chosen were highly-disparate in their null hypotheses and approaches to LD, making

it difficult to uniformly compare their results. Alternatively, multi-site collaborations might simultaneously analyze several large data sets using a small number of analytical tools in the same conceptual category; comparisons of the results would advance the underlying science and critically evaluate tools against closely-related options.

Finally, methods for integrating different types of association signals are developing. A nascent view proposes that combining genome-wide expression and genotyping data into a joint quantitative signal can increase power for discovery [61, 91, 136, 137]. One particularly attractive feature of this view is that it augments structure (genotype) with function (expression). Indeed, one study demonstrated that SNPs correlated with gene expression changes (expression quantitative trait loci; eQTLs) were more likely to show disease association than other SNPs from a GWAS array [138]. Relatedly, visualization tools can graphically overlay association metrics onto other data in order to prioritize markers. Visualization has been used to integrate SNP association with quantitative imaging phenotypes [139], among other examples.

H. Pathways and networks: bridging multi-omics data

As pathway analysis of genomic data has exploded in use, its methods have matured, its results are beginning to meet its potential, and points of consensus are emerging for its continued application and future development. In the coming years, we anticipate that pathways and networks will assume a farther-reaching role in view of the need to integrate multi-omics data through systems biology approaches [140, 141]. A variety of large-scale strategies are active in the study of complex diseases, including genomic, transcriptomic, proteomic, and metabolomic

approaches, and data from all of these sources can be analyzed through pathways and networks representing coordinated functions and relationships. Importantly, while gene associations do not always indicate therapeutic targets [42], pathways and networks implicated by analyses at multiple levels would be prime targets for therapies. Integrating large-scale data assayed through diverse strategies related to structure and function would provide a fertile process for exploring connections between replicable, statistical association and meaningful biology. As such, the role of pathways and networks as the hub for this integration will be vital in the years to come.

III. Genome-wide pathway analysis of episodic memory performance in the Alzheimer's Disease Neuroimaging Initiative (ADNI) cohort

A. Introduction

Human memory is a complex, dynamic trait with significant roles in development, aging, and disease. Impairment in episodic memory – involving the encoding and conscious recollection of experiences – is an early hallmark feature of AD, the most common cause of dementia [7]. Declines in episodic and other memory domains are also found in normal cognitive aging and many age-related disorders, including Parkinson's disease (PD), diabetes, and cancer [142]. With the rising incidence and burdens of dementia, a better understanding of its causes is crucial for the development of memory-sparing lifestyle and drug therapies [4, 143].

At present, the molecular mechanisms underlying memory performance in AD and other clinical settings are not fully understood. Epidemiological studies have linked many factors to memory, including the presence of vascular and metabolic disease, mental and physical activity, and educational and occupational attainment [144-146]. Memory is also estimated to have substantial heritability (30-60%) based on twin studies and is thought to be influenced by common and rare genetic variation in multiple pathways [145]. Although GWAS and candidate gene studies have implicated numerous SNPs in memory performance [15, 145, 147-152], significant heritability – and biological understanding – remains unexplained.

An important consideration for addressing this knowledge gap is that genetic studies of quantitative endophenotypes [11] such as memory performance are highly dependent on the

quality of the phenotypic data. There are numerous extant metrics for assessing memory in amnestic populations, and these metrics can have differential sensitivities to deficits in various memory domains and sub-domains [153]. As a result, studies often attempt to leverage the relative strengths and weaknesses of these metrics by creating composite scores from multiple assessments given to participants [77, 154]. Recent advances in psychometric approaches, which attempt to empirically generate and validate composite scores based on item-level responses from multiple memory instruments, may yield optimized measures of cognitive functioning that provide greater power for genetic studies [155-158].

The complex etiology of memory performance adds another challenge for genetic studies. Human and animal model investigations have demonstrated that the complex processes of memory consolidation and recall involve numerous and diverse cellular and molecular pathways [142]. While GWAS of complex phenotypes have historically focused on identifying individual susceptibility loci, their efficacy has been confounded by several factors. Most common alleles implicated by GWAS have exhibited modest effect sizes [159]. In addition, robust genetic associations have not always served as appropriate therapeutic targets [42]. Further, it is well-understood that genes do not exist in isolation, but instead function as sets within biological pathways and networks [31, 32, 49]. As a result, GWAS of complex phenotypes are increasingly being analyzed through statistical methods designed to identify biological pathways enriched with association to those phenotypes [31, 32, 41].

Although genome-wide pathway analysis has been performed for complex neurological phenotypes, including brain glutamate levels [117], cerebrospinal fluid (CSF) Aβ levels [160], and information processing speed [161], this strategy has not been previously applied to memory

performance. Here, we perform the first genome-wide pathway analysis of memory performance, using a psychometrically-derived episodic memory score for participants from the Alzheimer's Disease Neuroimaging Initiative (ADNI). Our analyses identify key pathways associated with memory and highlight highly-represented genes in these pathways as key targets for future studies. We also use network analysis and data from a public human brain tissue expression database to isolate sets of genes which are co-regulated and/or co-expressed. The enriched pathways and gene networks identified in this analysis suggest prime targets for further studies of AD, memory impairment, and normal cognitive aging, and further demonstrate the efficacy of pathway-based approaches for analyzing GWAS of complex phenotypes.

B. Methods

Study participants

This study utilized data from the initial phase of ADNI [162] (http://adni.loni.ucla.edu/), a multi-site longitudinal study that was launched in 2004 as a public-private partnership. The initial phase of ADNI enrolled individuals aged 55-90 years who were recruited from over 50 sites across the United States and Canada and followed at 6- to 12-month intervals for 2-3 years. These individuals included approximately 200 healthy controls (HC), 400 patients with late MCI (LMCI), and 200 patients clinically diagnosed with probable AD. As described elsewhere [162-164], diagnoses of participants were made on a clinical basis (via neuropsychological assessment data and patient and informant reports of cognitive performance and functioning in activities of daily living) at consensus conferences involving neurologists, neuropsychologists, and study

coordinators. All participants provided written informed consent and study protocols were approved by each site's institutional review board. Further information about ADNI, including full study protocols, complete inclusion and exclusion criteria, and data collection and availability can be found at http://www.adni-info.org/.

Sample characteristics across diagnostic groups were evaluated using IBM SPSS 19.0. A one-way analysis of variance was performed for continuous variables and a Pearson chi-square test was performed for categorical variables.

Psychometrically-derived composite episodic memory scores

All participants (original *N* = 818) were administered an extensive neuropsychological assessment, including several measures of memory, at each study visit. For each subject, a composite score for episodic memory at the baseline visit was calculated as described previously [155] through analysis of item-level data from the ADNI neuropsychological battery. Briefly, the authors used psychometric theory to select test battery items which could be considered as indicators of episodic memory functioning. An iterative process of confirmatory factor analysis was used to construct the final, optimized model for describing episodic memory performance at baseline. In particular, the following item-level tests were applied to the final model: the memory sub-scores from the Mini-Mental Status Examination [165, 166]; the immediate and delayed recall and recognition scores on a word list learning task from the Alzheimer's Disease Assessment Scale-cognitive subscale [167]; all immediate and delayed recall and recognition scores from the Rey Auditory Verbal Learning Test [168]; and all immediate and delayed recall scores on Logical Memory prose passages from the Wechsler Memory Scale-Revised [169]. The

final model exhibited excellent fit based on standard criteria (Confirmatory Fit Index > 0.95, Tucker Lewis Index > 0.95, and Root Mean Squared Error of Approximation < 0.05) [158]. A composite episodic memory score could not be calculated for eight participants due to incomplete item-level data at baseline.

Genotyping and quality control

Details on genotyping for the ADNI sample have been described previously [170]. All participants in this study were genotyped according to the manufacturer's protocol using the Human610-Quad BeadChip (Illumina, Inc., San Diego, CA), which included 620,901 SNPs and structural variant markers. In addition, given the strong association of *APOE* with MCI and AD susceptibility [14, 171, 172], the SNPs (rs429358, rs7412) that characterize the *APOE* ε2, ε3, and ε4 alleles were genotyped separately due to not being available on the GWAS array [170, 173].

Genotype data (original N = 818) was subjected to stringent quality control procedures as described previously using PLINK, version 1.07 [94, 174]. SNPs were excluded if they had a call rate < 90%, Hardy-Weinberg equilibrium test $p < 10^{-6}$, or minor allele frequency (MAF) < 5%. Samples were excluded if they had a call rate < 90% (1 participant), ambiguous gender identification (2 participants), or failed an identity check (3 participants). To limit the possible effects of population substructure, analyses were restricted to participants with non-Hispanic Caucasian ancestry determined by multidimensional scaling analysis as described previously [175]; this resulted in the exclusion of 62 participants. Following all quality control procedures, 750 participants and 531,096 SNPs were designated for subsequent analyses and the genotyping rate was > 0.995 among the remaining samples.

GWAS of episodic memory performance

Of the 750 participants with quality controlled genotype data, 742 participants were identified as having a psychometrically-derived composite episodic memory score at baseline. A GWAS of this composite memory score was performed using linear regression under an additive genetic in PLINK. Demographic factors with known influences on memory or cognition were included as covariates, including age at the baseline visit, education, gender, and handedness. The direct and inverse relationships, respectively, of age and education level on memory decline have been well-established [176], while putative effects of gender and handedness on cognition are subjects of active exploration [177, 178] and were included as part of a conservative approach. *APOE* ε4 allele status (presence vs. absence) was also used as a covariate in the GWAS to account for the largest known genetic influence on memory performance in an MCI- and AD-based clinical population [179] in order to focus on identification of novel biological influences. For all SNPs included in the analysis, a *p*-value was generated to represent the nominal association of that SNP to the composite memory score. Manhattan and Quantile-Quantile (Q-Q) plots for the GWAS were generated using PLINK and Haploview [180].

Pathway analysis

All SNPs included in the GWAS were mapped to genes using the NCBI Build 36.1 reference sequence [181]. An extended gene mapping window of ±20 kb was used to account for SNPs belonging to putative regulatory regions; this resulted in some SNPs being mapped to more than one gene. In total, 277,615 SNPs were assigned to 17,456 genes. Pathway annotations, representing gene sets defined by membership in biological pathways, were downloaded from

the Molecular Signatures Database [109], version 3.0. This annotation data comprised a collection of canonical, expertly-curated pathways from three publically-available databases, BioCarta (http://www.biocarta.com/), the Kyoto Encyclopedia of Genes and Genomes (KEGG; http://www.genome.jp/kegg/), and Reactome (http://www.reactome.org/). In total, 818 pathways were downloaded, included 217 pathways from BioCarta, 186 pathways from KEGG, and 430 pathways from Reactome.

The GSA-SNP software [108] was used to assess for pathways enriched against the composite memory score. This software uses a competitive enrichment algorithm [98], where the null hypothesis holds that a pathway-phenotype association is not significantly different from all other pathway-phenotype associations under analysis. Competitive enrichment strategies are robust to the effects of genomic inflation due to population stratification or other confounding factors [51, 182]. In GSA-SNP, the significance score for each gene under analysis was calculated as the -log of the *k*-th best SNP-level *p*-value in the gene. Corresponding with the authors' recommendation [108], we selected $k = 2$ to limit the effects of both single, highly-significant loci and of spurious SNP-level associations on driving pathway enrichment. Each pathway was then assessed for phenotype enrichment by the *Z*-statistic method [183], which incorporates the gene-wide significance scores and the number of genes within each set.

Since small pathways can exhibit spurious phenotype associations due to large single locus effects [51], and since large pathways are more likely to exhibit association by chance alone [79], analysis was restricted to the 280 pathways containing 10-200 genes. To correct for multiple hypothesis testing, the False Discovery Rate (FDR) [124] was applied to the *p*-values generated by the enrichment algorithm. For pathways enriched at an FDR-corrected *p*-value <

0.05, we analyzed their constituent genes to obtain a count of each gene's occurrences in those enriched pathways. Genes that were highly-represented among the enriched pathways (defined as being constituents of > 15% of the enriched pathways) were isolated for follow-up analyses.

Transcription factor network analysis

We further investigated the list of highly-represented genes from our enriched pathways through network analysis using MetaCore (GeneGo, Inc.). In particular, we applied the transcription factor network analysis algorithm to identify subsets of those genes with coordinate regulation by known transcription factors. *APOE* was also included in these analyses, given its well-characterized association with MCI and AD and their related memory deficits.

Gene expression analysis using the Allen Human Brain Atlas

We also interrogated the list of highly-represented genes from our enriched pathways for their expression profiles in normal brain tissue using the Allen Human Brain Atlas (Allen Institute for Brain Science, Seattle, WA; available from http://www.brain-map.org/). The Allen Human Brain Atlas includes genome-wide microarray-based expression profiles in postmortem brain tissue from subjects with no known neuropsychiatric or neuropathological history. These expression profiles cover the entire brain through systematic sampling of regional tissue, and are integrated with multi-modal brain imaging and other data for visualization and analysis. Detailed information on tis database is available on-line (http://human.brain-map.org/docs.html/). We employed the Allen Human Brain Atlas to examine genes of interest (including highly-represented genes from our enriched pathways) for patterns in their expression profiles. In

particular, we used the heat map tool (which visually displays normalized expression for a gene probe across 25 large neuroanatomic regions) and correlational analysis (which calculates a Pearson correlation coefficient between the expression profiles of two gene probes) to identify a set of key genes with high ($r > 0.7$) co-localization and co-expression.

C. Results

Demographic characteristics and mean composite memory scores for all diagnostic groups (HC, LMCI, and AD) are presented in Table 3. While baseline age and handedness were not significantly different across diagnostic groups, gender exhibited a significant difference ($p < 0.05$), with males relatively overrepresented among LMCI subjects. In addition, as expected, education level and *APOE* ε4 allele status exhibited significant differences across groups ($p < 0.001$). Also as expected, composite memory scores differed across all diagnostic groups, including all pairwise group comparisons ($p < 0.001$).

In order to assess SNP associations to the composite memory scores in this sample, we performed a GWAS with the addition of five covariates. The GWAS failed to identify any SNPs with significant association to the composite memory score at a Bonferroni-determined threshold *p*-value of 9.42×10^{-8} (i.e., 0.05/531,096). The peak SNP in the association analysis was rs9890008 (Chr 17), which has not been mapped to a known gene and which exhibited an unadjusted *p*-value of 2.21×10^{-6}. Overall, 25,960 SNPs showed nominal *p*-values < 0.05 (unadjusted). Manhattan (Figure 4) and Q-Q (Figure 5) plots are displayed for the GWAS.

Table 3. Selected characteristics for the study sample. Values represented are mean ± SD unless specified otherwise.

	HC *(N = 207)*	**LMCI** *(N = 362)*	**AD** *(N* **= 173)**	***p*-value**[a]
Age at baseline	76.1 ± 5.0	74.9 ± 7.4	75.6 ± 7.5	0.139
Gender (male/female)	113/94	234/128	92/81	0.012
Years of education	16.2 ± 2.7	15.7 ± 3.0	14.9 ± 3.0	< 0.001
Handedness (right/left)	191/16	328/34	161/12	0.586
***APOE* ε4 allele (absent/present)**	152/55	165/197	61/112	< 0.001
Composite memory score[b]	1.05 ± 0.59	-0.14 ± 0.64	-1.01 ± 0.63	< 0.001

[a]For categorical variables, *p*-values were computed using the Pearson chi-square tests; for continuous variables, *p*-values were computed using a one-way analysis of variance.

[b]Scores for individual participants were represented as z-scores with a defined mean of 0, and standard deviation of 1, based on the 810 participants with complete item-level memory data at baseline.

Figure 4. Manhattan plot for the GWAS of episodic memory in ADNI. The x-axis refers to positions along the genome (separated by chromosome) for each SNP (represented by a dot) included in the analysis. The y-axis refers to the negative logarithm of the *p*-value for the test of association between each SNP and the quantitative memory phenotype. No SNPs exhibited genome-wide significant association (red line) to the composite episodic memory score, while 40 SNPs exhibited suggestive association ($p < 5 \times 10^{-5}$, blue line). The 5 most significant, independent ($r^2 < 0.2$) SNPs are labeled along with their corresponding genes (if known).

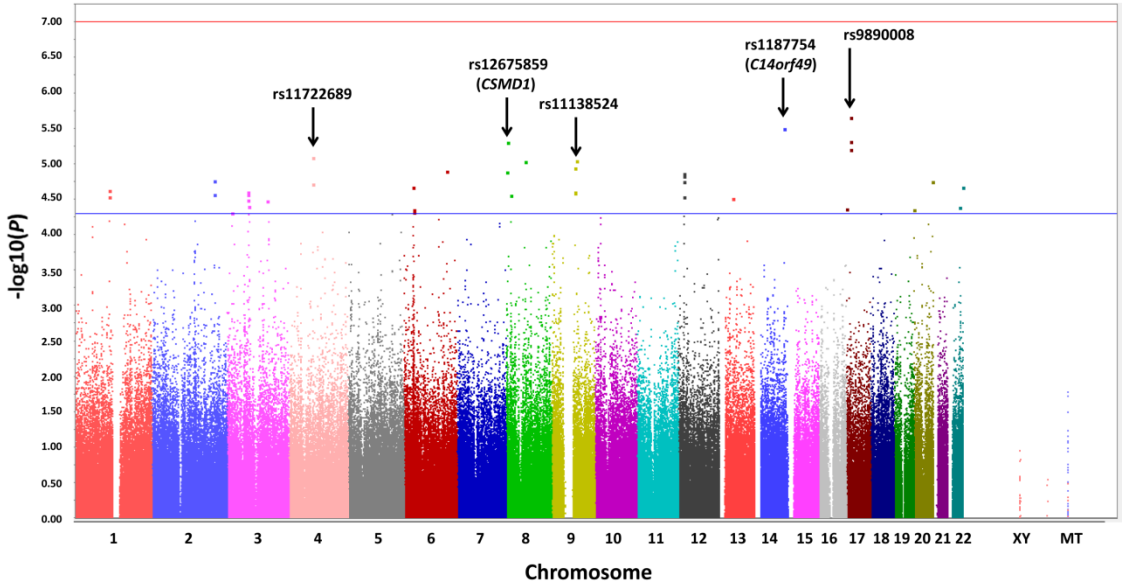
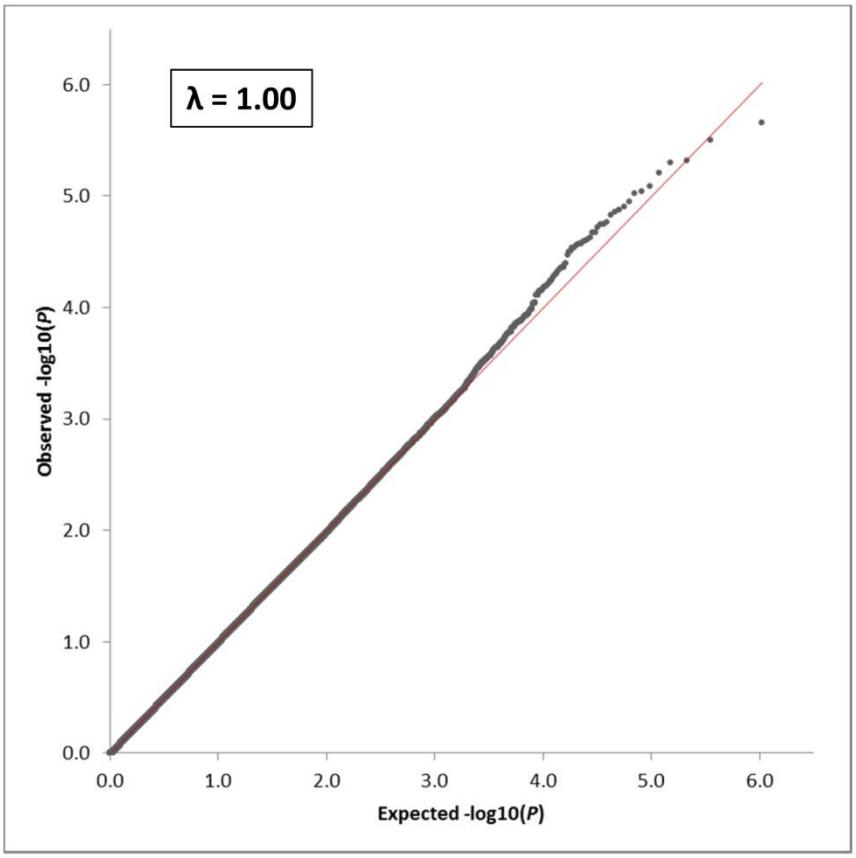


Figure 5. Q-Q plot for the GWAS of episodic memory in ADNI. No evidence of spurious inflation of association tests statistics was identified.



Notably, when *APOE* ε4 allele status was removed as a covariate, one SNP (rs2075650) residing in the *TOMM40* (translocase of outer mitochondrial membrane 40) gene exhibited genome-wide significant association ($p = 2.19 \times 10^{-9}$) to episodic memory. This result was probable, given that the *TOMM40* gene is adjacent to and often considered as one locus with *APOE* on chromosome 19. In addition, *TOMM40* variants have been associated with late-onset AD [173, 184, 185] as well as with structural brain and cognitive function changes suggestive of presymptomatic late-onset AD [186].

Next, the *p*-value output from the GWA analysis was used as input for pathway enrichment analysis. Using the GSA-SNP software tool, we identified 27 canonical pathways with enrichment (FDR-corrected p-value < 0.05) against the composite memory score (Table 4). Following these analyses, we examined the enriched pathways in detail in order to better characterize their biological import in memory deficits. First, we used existing knowledge and insight to conceptually categorize the 27 enriched pathways into 4 broader realms of biology (Figure 6). In particular, 11 enriched pathways represented classical cellular and molecular processes essential in normal memory consolidation signaling [142]. These pathways included functions of neurotransmitter receptor activation, downstream calcium-mediated signaling, and long-lasting potentiation of synaptic strength, among other processes. In addition, six pathways related to cell adhesion were enriched, including focal adhesion pathways from both the Reactome and KEGG databases, and interactions involving neuronal cell adhesion molecule 1 (NCAM1). Finally, four enriched pathways were related to neuronal differentiation and guided axonal growth, while a further six enriched pathways were involved in inflammation or other complex signaling processes. Notably, while we restricted analysis to pathways containing 10-200 genes, the enrichment results were nearly identical when upper limits of 300 or 400 genes

Table 4. Pathways showing enrichment of association to episodic memory in ADNI.

Pathway (Gene Set) Name	Set Size[a]	*p*-value	FDR *p*
[b]Transmission across chemical synapses	136 (122)	2.14×10^{-7}	1.77×10^{-4}
[c]Calcium signaling pathway	184 (165)	2.82×10^{-7}	1.17×10^{-4}
[c]Type I diabetes mellitus	50 (41)	7.90×10^{-6}	0.002
[b]Neurotransmitter receptor binding and downstream transmission	90 (78)	2.84×10^{-5}	0.006
[c]Arrhythmogenic right ventricular cardiomyopathy	82 (72)	3.11×10^{-5}	0.005
[b]SLC-mediated membrane transport	175 (162)	3.87×10^{-5}	0.005
[c]Focal adhesion	207 (187)	4.68×10^{-5}	0.006
[c]Axon guidance	135 (123)	4.68×10^{-5}	0.005
[c]Long-term depression	76 (65)	9.02×10^{-5}	0.008
[b]Axon guidance	167 (154)	9.58×10^{-5}	0.008
[c]Adherens junction	81 (71)	1.04×10^{-4}	0.008
[b]Other semaphorin interactions	22 (15)	1.20×10^{-4}	0.008
[b]NCAM1 interactions	50 (42)	1.37×10^{-4}	0.009
[c]Long-term potentiation	76 (65)	3.14×10^{-4}	0.019
[b]Activation of glutamate NMDA receptor and post-synaptic events	42 (32)	5.02×10^{-4}	0.028
[c]Cell adhesion molecules (CAMs)	140 (123)	5.60×10^{-4}	0.029
[b]SEMA3A plexin repulsion signaling by inhibiting integrin adhesion	20 (13)	6.06×10^{-4}	0.030
[c]Tryptophan metabolism	46 (35)	6.53×10^{-4}	0.030

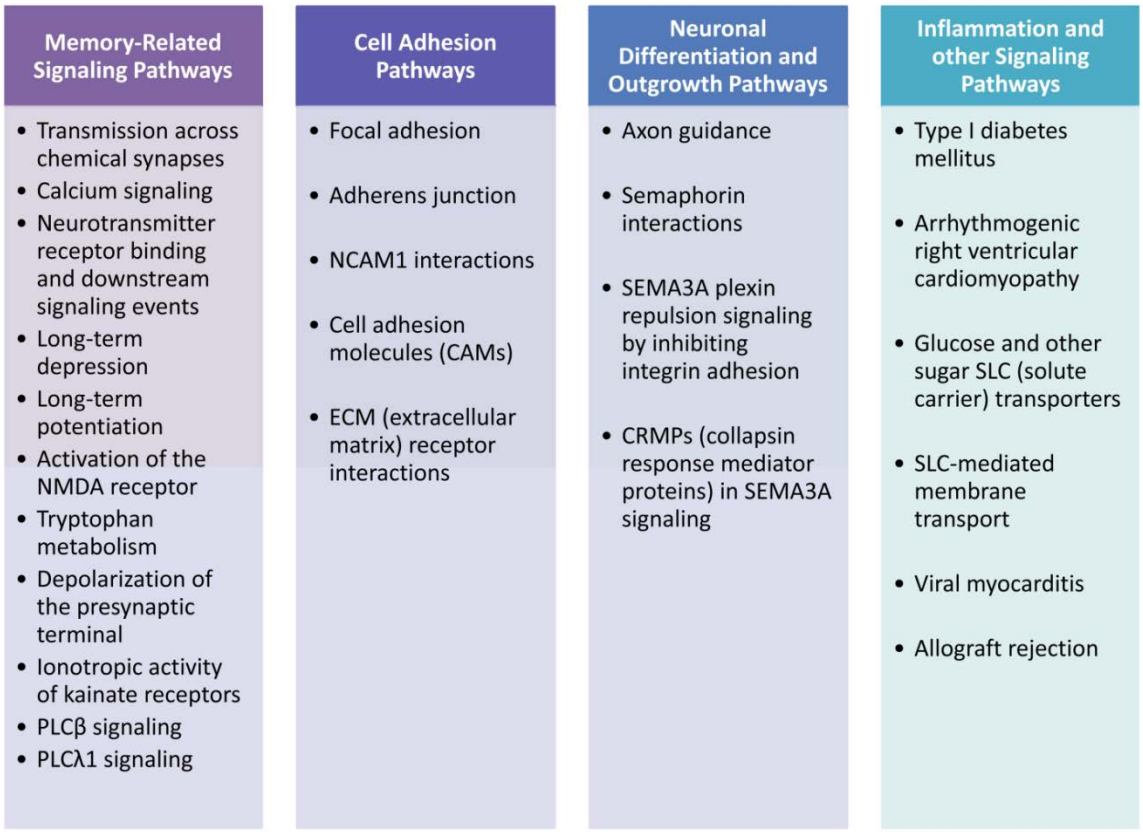
[b]Depolarization of the presynaptic terminal triggers the opening of calcium channels	18 (12)	6.71×10^{-4}	0.029
[b]PLCβ-mediated events	44 (35)	6.91×10^{-4}	0.029
[c]Viral myocarditis	79 (67)	9.98×10^{-4}	0.039
[c]Allograft rejection	44 (34)	0.001	0.039
[b]Glucose and other sugar SLC transporters	88 (80)	0.001	0.039
[b]Ionotropic activity of kainate receptors	18 (11)	0.001	0.042
[c]ECM receptor interaction	90 (83)	0.001	0.041
[b]PLCγ1 signaling	41 (32)	0.001	0.044
[b]CRMPs in SEMA3A signaling	22 (15)	0.001	0.045

[a]Entries are displayed as: number of genes in the set (number of genes from the GWA data)

[b]Reactome pathway

[c]KEGG pathway

Figure 6. Conceptual classification of pathways showing enrichment of association to episodic memory in ADNI.



were used: in those cases, two additional Reactome pathways exhibited enrichment (transmembrane transport of small molecules, FDR-corrected *p*-value = 0.029; adherens junction interactions, FDR-corrected *p*-value = 0.049).

As further follow-up, we identified 44 genes that were highly-represented across the 27 enriched pathways (Table 5). Half (22) of these 44 genes were constituents of 6 or more enriched pathways, suggesting that variants in those genes can have wide-ranging roles in mediating memory impairment due to their diverse functions. We also assessed for underlying transcriptional relationships among these highly-represented genes and *APOE*, given the latter's singular association with MCI and AD and their related memory deficits. Using network analysis in MetaCore, we discovered that 14 of the 22 most-represented genes from our analyses were part of a transcriptional regulation network driven by the specificity protein 1 (SP1) transcription factor and involving *APOE* and the *APOE* receptor-2 (Figure 7).

Finally, we used data from the Allen Human Brain Atlas to evaluate if the identified genes of interest exhibited co-expression in normal brain tissue. Through heat map visualization and correlational analysis, we identified a set of 10 key genes with strong co-expression (Pearson $r > 0.7$) across the major neuroanatomic regions of the brain (Figure 8a). In particular, 6 of these genes (*CAMK2A*, *CACNB1*, *CALM1*, *CALM3*, *GRIN2A*, and *MAPK1*) were highly-represented among our enriched pathways, while the other 4 genes (*CDK5*, *GSK3B*, *GRIN2B*, and *PRNP*) were constituents of our enriched pathways that were also known AD susceptibility genes found in the AlzGene database (http://www.alzgene.org/) [127]. An example of the cortical and subcortical expression patterns common to this gene set was also generated for one of these

Table 5. Highly-represented genes among the pathways showing enrichment of association to episodic memory in ADNI.

Occurrences	Gene ID	Gene Name
9	*MAPK1*	mitogen-activated protein kinase 1
8	*CALM1*	calmodulin 1 (phosphorylase kinase, delta)
8	*CALM2*	calmodulin 2 (phosphorylase kinase, delta)
8	*CALM3*	calmodulin 3 (phosphorylase kinase, delta)
8	*HRAS*	v-Ha-ras Harvey rat sarcoma viral oncogene homolog
7	*ADCY1*	adenylate cyclase 1 (brain)
7	*ADCY8*	adenylate cyclase 8 (brain)
7	*CAMK4*	calcium/calmodulin-dependent protein kinase IV
7	*FYN*	FYN oncogene related to SRC, FGR, YES
7	*ITGB1*	integrin, beta 1 (fibronectin receptor, beta polypeptide, antigen CD29)
7	*PRKACB*	protein kinase A, cAMP-dependent, catalytic, beta
7	*RAF1*	v-raf-1 murine leukemia viral oncogene homolog 1
6	*ACTN2*	actinin, alpha 2
6	*ADCY3*	adenylate cyclase 3
6	*CREB1*	cAMP responsive element binding protein 1
6	*MAPK3*	mitogen-activated protein kinase 3
6	*PLCB1*	phospholipase C, beta 1 (phosphoinositide-specific)
6	*PLCB2*	phospholipase C, beta 2

6	*PLCB3*	phospholipase C, beta 3 (phosphatidylinositol-specific)
6	*PRKCA*	protein kinase C, alpha
6	*PRKCG*	protein kinase C, gamma
6	*RAC1*	ras-related C3 botulinum toxin substrate 1 (rho family, GTP binding protein)
5	*CACNA1C*	calcium channel, voltage-dependent, L type, alpha 1C subunit
5	*CACNB1*	calcium channel, voltage-dependent, beta 1 subunit
5	*CACNB2*	calcium channel, voltage-dependent, beta 2 subunit
5	*CACNB3*	calcium channel, voltage-dependent, beta 3 subunit
5	*CACNB4*	calcium channel, voltage-dependent, beta 4 subunit
5	*CAMK2A*	calcium/calmodulin-dependent protein kinase II alpha
5	*CAMK2B*	calcium/calmodulin-dependent protein kinase II beta
5	*CAMK2D*	calcium/calmodulin-dependent protein kinase II delta
5	*CAMK2G*	calcium/calmodulin-dependent protein kinase II gamma
5	*GRIA1*	glutamate receptor, ionotropic, AMPA 1
5	*GRIA2*	glutamate receptor, ionotropic, AMPA 2
5	*GRIN1*	glutamate receptor, ionotropic, N-methyl D-aspartate 1
5	*GRIN2A*	glutamate receptor, ionotropic, N-methyl D-aspartate 2A
5	*GRIN2C*	glutamate receptor, ionotropic, N-methyl D-aspartate 2C
5	*GRIN2D*	glutamate receptor, ionotropic, N-methyl D-aspartate 2D
5	*ITGA1*	integrin, alpha 1
5	*ITPR1*	inositol 1,4,5-trisphosphate receptor, type 1
5	*ITPR2*	inositol 1,4,5-trisphosphate receptor, type 2

5	*ITPR3*	inositol 1,4,5-trisphosphate receptor, type 3
5	*PLXNA1*	plexin A1
5	*PLXNA2*	plexin A2
5	*RAC2*	ras-related C3 botulinum toxin substrate 2 (rho family, GTP binding protein)

Figure 7. Transcriptional regulatory network centered on the SP1 transcription factor involves many genes of interest from the pathway analysis of episodic memory in ADNI. Green arrows indicate positive regulatory effects, red arrows indicate negative regulatory effects, and gray arrows indicate unspecified regulatory effects. The primary image was generated through the MetaCore software.

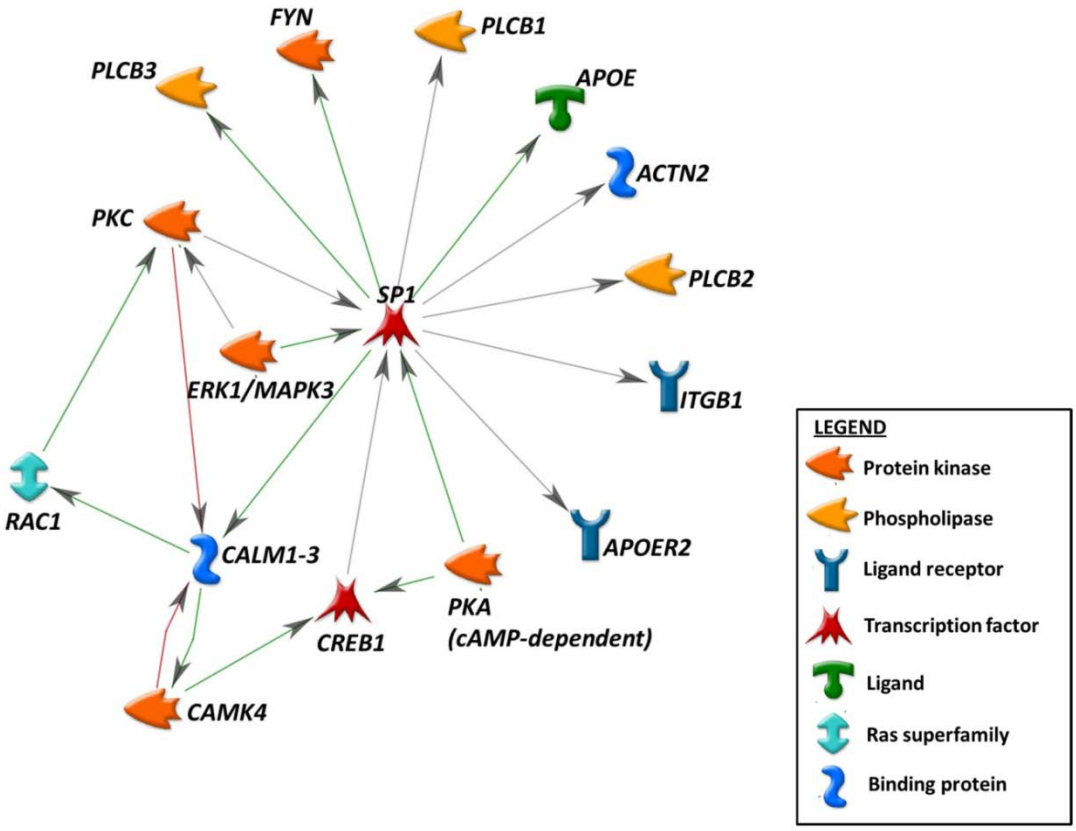
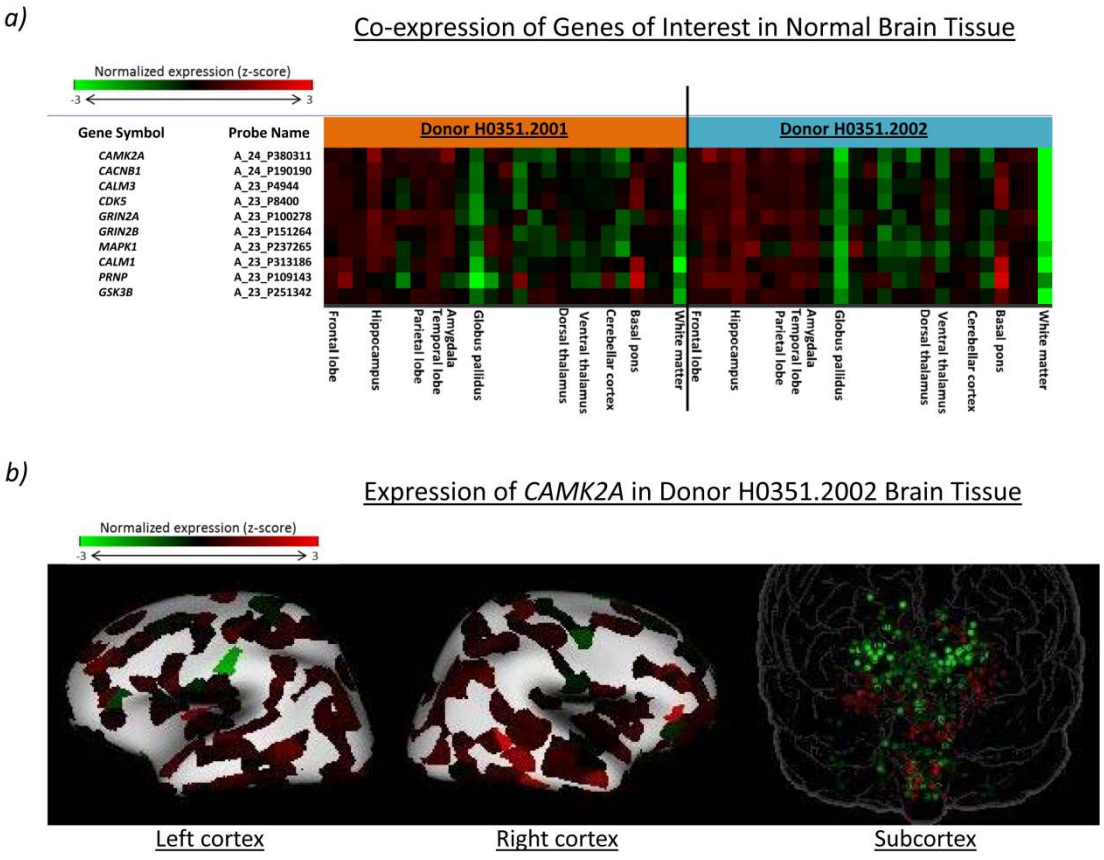


Figure 8. Genes of interest for episodic memory performance exhibit co-expression in postmortem human brain tissue from healthy subjects in the Allen Human Brain Atlas. (A) Normalized, microarray-based expression profiles across 25 major neuroanatomic regions of the brain are provided for 10 key genes of interest. Moving from left to right on the heat map is analogous to moving from anterior to posterior regions first in the cortex, followed by subcortical areas and then the cerebellum and brainstem. The genes represented exhibit strong co-expression (Pearson $r > 0.7$) across the brain in data from two subjects. (B) For the co-expressed gene set, a representative cortical and subcortical expression profile is shown for the *CAMK2A* gene.



genes, *CAMK2A* (Figure 8b). These findings suggest further underlying functional relationships between MCI and AD disease risk and the pathogenesis of memory impairment.

D. Discussion

In this study, we used a psychometrically-optimized composite measure of episodic memory performance as a phenotype for GWAS in a sample of controls plus LMCI and AD patients. Through a genome-wide pathway analysis, we identified 27 canonical pathways showing enrichment of association to episodic memory performance in this sample. These enriched pathways suggest that the genetic architecture of memory impairment in this clinical sample of the AD spectrum spans both processes classically understood to be involved in normal memory consolidation as well as mechanisms with broader roles in cognition and aging, such as those involving neuronal cell adhesion and inflammation.

It should be emphasized that pathway-based approaches analyze genetic data in the context of operative functional groups; as a result, pathway analysis findings are uniquely and naturally connected to the functional biology underlying complex phenotypes. This insight is vital for future investigations, given that pathway mechanisms are principal sources for developing strategies to diagnose, treat, and prevent complex disorders. It is also important to note that our analysis elucidated pathways with robust enrichment despite using GWAS input data that included a relatively modest distribution of SNP-level phenotype associations. These results affirm the potential of pathway-based analytical approaches to detect significant relationships that are otherwise concealed within single-SNP or single-gene analysis. The use of genome-wide pathway analysis in this study also facilitated the detection of unexpected relationships with

memory performance, including pathways not classically related to memory signaling, and subsequently, interesting transcriptional and expression networks. While targeted "candidate pathway" approaches have advantages, these unexpected relationships would not have been easily predicted as candidates for analysis based on prior knowledge.

At a functional level, the enriched pathways identified in this study present interesting biological implications in relation to memory impairment. In a sense, it might be expected that cellular and molecular processes classically understood as mediating memory consolidation would constitute a major part of the genetic architecture of memory impairment. However, the processes underlying memory consolidation are numerous and diverse, and to date it has not been clear which specific pathways are essential objects of the impact of genetic variants. Our pathway enrichment study highlights potential major components of this genetic architecture.

In particular, we observed significant enrichment of pathways related to neurotransmitter receptor activation and downstream signaling. These pathways and their resultant calcium-mediated signaling are vital in converting short-term memories, which exist as axonal firing patterns, into long-lasting changes in synaptic strength [187]. As a parallel, it makes sense that a composite long-term potentiation pathway (comprising multiple processes leading to long-lasting increases in synaptic strength) would include genetic determinants related to memory performance. Our results also indicate the need for further exploration of long-term depression as a substrate for memory impairment, particularly given its proposed roles in mediating the cognitive effects of acute stress and synaptic pruning in neurodegenerative diseases [188].

Other pathways with enrichment of association to memory in this study have previously been implicated in neuronal development and cognition. For example, neuronal cell adhesion molecules (NCAMs) appear to play major roles in susceptibility for schizophrenia, bipolar disorder, and autism-spectrum disorders [189]. Genetic variants of NCAMs have also been associated with CSF biomarkers for AD [160], and cell adhesion molecule pathways have exhibited enrichment in a genome-wide pathway analysis of AD case-control status [59]. In addition, expression of NCAMs in cholinergic neurons appears to increase susceptibility to AD-related neurodegeneration [190], and there is emerging evidence of interactions among NCAMs, the MAPK pathway, and Aβ precursor protein [191]. More broadly, these findings suggest a prominent role for cell adhesion pathways in maintaining the processes of synaptic plasticity that are believed to underlie learning and memory [192].

It is also interesting that pathways on axon guidance, including those involving functions of ephrins, semaphorins, and rho GTPases, were enriched in this study. Axon guidance pathways are key in forming guided neuronal network connections, and have been previously implicated in early neuronal development and associated genetic conditions [193]. Together with these new enrichment results, the proposed interaction between vascular and neuronal factors related to axon guidance [194] in relation to memory may be an important direction for further studies. In addition, given the complex interactions among brain cells and immune system functions, the immune-related pathways enriched in this study suggest additional candidates for modulation of memory and synaptic plasticity [195]. It may be particularly fruitful to examine immune mediators of memory dysfunction that exert influences independent from Aβ-related activation of microglia in AD [196].

A related perspective – and additional interesting targets for future investigations – can be achieved by examining the set of genes that were highly-represented across the enriched pathways in this study. Prominent groups of gene products represented in this set are particularly important in memory consolidation. For example, integrins, cadherins, and alpha-actinin are known to regulate neuronal cytoskeletal structure to mediate synaptic plasticity and are proposed to signal through MAPK cascades for localized protein synthesis at the specific dendrites being activated to precisely potentiate their synaptic connections [142]. Another important group of gene products is related to the calcium influx that follows neurotransmitter receptor activation at synapses: this calcium influx leads to activation of a signaling axis involving calmodulin, protein kinases (PKA, PKC-α, CAMKII subtypes, and CAMKIV), and transcription factors (CREB subtypes), among other molecules [142].

Overall, since the highly-represented genes from our data act in numerous pathways, our results reinforce the benefits of studying genetic variation within a pathway-based framework: in this context, variants of moderate individual effect sizes can nevertheless be identified as exerting strong and wide-ranging effects when juxtaposed with other meaningful variants in shared functional processes [31]. Extensions of this pathway-based analytic framework will be extremely valuable in identifying localized effects of specific pathways on particular brain regions, particularly given that imaging correlates have been identified for loci with known effects on memory, such as the impact of *WWC1/KIBRA* gene variants on hippocampal activation [197]. Notably, innovative voxelwise SNP- [198, 199] and gene-based [200] imaging genetics approaches have been successfully employed in studies of AD, as has a novel method for generating multivariate “genetic components” for imaging analysis [201]. These strategies

will serve as rich foundations for future pathway-based imaging genetics analyses to complement those focusing on cognitive performance.

In addition, the network analyses in this study reinforce the notion that key genes related to memory impairment function in coordination. We found that a preponderance of the most highly-represented genes in our enriched pathways were constituents of a transcriptional regulation network driven by the SP1 transcription factor (Figure 4). The SP1 transcription factor has known binding regions in the promoters of genes related to Aβ precursor protein [202, 203], tau protein [204], and *APOE*. In particular, SP1 has been proposed as a regulator of *APOE* promoter activity in relation to two promoter polymorphisms with significant association to AD [205]. Given that networks of common regulation represent prime targets for identifying common functions, further investigation of the transcriptional network that we have identified may elucidate the as-yet-unknown mechanistic connections among *APOE* and other susceptibility loci, AD pathogenesis, and MCI- and AD-related memory impairment.

Finally, expression analysis using the Allen Human Brain Atlas revealed additional functional relationships among key genes. Since strong co-expression of a set of 10 key genes in the brain may indicate common modes of function, further study of this and other similar sets may be of great value. In addition, the co-expression of highly-represented genes from the enriched pathways in this study with known AD susceptibility genes suggests the possibility of significant crosstalk between AD pathogenesis and basic memory processes. While the data in the Allen Human Brain Atlas has several limitations, including a small number of subjects and the inclusion of only postmortem brain tissue from neuropsychologically- and neuropathologically-normal subjects, at present it is the only available resource which integrates multi-modal brain imaging

data with whole- and regional-brain genome-wide expression data. As such, this and other functional annotation resources will be vital for identifying mechanistic connections between AD pathogenesis and memory impairment, including future efforts to quantitatively assess the significance of overlap between memory pathways and AD pathways.

There are some notable limitations to the current study. First, a pathway analysis is only as good as the functional information underlying its pathway definitions. Importantly, some intragenic SNPs may not affect the function or expression of their assigned gene, while other SNPs may functionally impact distant genes or even multiple genes [89, 90]. As functional annotation of the genome becomes more extensive, the power of pathway analyses will heighten. For this study, we used a collection of canonical pathways curated through expert review. While these pathway annotations are expected to have high accuracy, differences across pathway databases can lead to divergent enrichment analysis results [79]. For example, similarly-named pathways can have vastly different gene constituents, while distinctly-named pathways can nevertheless include significant gene overlap. As a result, an early discussion of pathway analysis methods recommended the use of multiple databases for each analysis [30]. While we have followed this recommendation for this analysis, future studies may benefit from formally assessing the relationships in biological coverage among the diverse pathways tested.

In addition, at this time there is no gold standard for pathway-based study design. Indeed, different enrichment algorithms and different parameters, such as those guiding SNP-to-gene mapping, can impact analytical results [135]. As such, pathway enrichment results benefit from further study using independent replication data sets and using alternative enrichment strategies. While differences across annotation resources, data sets, and analytical strategies

may impact SNP- or gene-level statistics [126], legitimately-associated pathways will likely exhibit significant enrichment or strongly-trending signal across a healthy percentage of studies. Finally, while it is beyond the scope of this study, future efforts will benefit from examining key memory-implicated genes and gene sets for epistatic (gene-gene) interactions with each other and with *APOE*.

There are also several caveats about the clinical setting for this study. First, the ADNI cohort represents a sample typical of a clinical trial for AD and MCI and is not a sample of the general population. As a result, the extent to which the present findings can be extended to account for episodic memory impairment in the general population remains to be determined. In addition, it is probable that the memory deficits in this study's MCI and AD participants are at least partially driven by AD-related pathology. While using *APOE* ε4 allele status as a covariate in these analyses likely attenuated this effect, a better understanding of the pathways underlying normal memory and other pathologies than AD may be achieved through studies of normal cohorts and other memory-impaired populations without AD-related pathology. In particular, further exploration of the relationships among *APOE* genotype status, Aβ load and pathology, and cognition in normal adults [206, 207] may be especially fruitful. Additionally, meta-analytic approaches to achieve larger study sample sizes may reveal greater SNP-level phenotype associations which could impact the pathway enrichment results. Finally, while this study used a composite episodic memory score optimized on the basis of modern psychometric theory, similar pathway-based studies using other quantitative memory phenotypes may provide different sensitivity and specificity to fine-grained memory deficits and would potentially serve as a validation for the discoveries of pathways enriched against the phenotype used in this study.

Nevertheless, the present results provide several new insights into key functional pathways associated with memory deficits in older adults with MCI or AD and controls. Importantly, these results highlight numerous candidates for further explorations of the SNPs, genes, and gene sets underlying normal memory processes and memory impairment. Overall, these findings encourage further use of pathway-based genetic analyses of quantitative memory phenotypes as statistically-powerful vehicles for discovery and as bridges to underlying biological mechanisms.

IV. Genetic modulators of cerebral amyloid deposition: a florbetapir PET GWAS and pathway analysis

A. Introduction

Cortical deposition of Aβ peptide is thought to be a crucial early step in the cascade of events leading to AD [208]. The presence of cortical neuritic plaques, consisting of Aβ fibrils surrounded by degenerating neuronal processes, is a hallmark feature for pathologic diagnosis of AD [209]. Aβ plaques have also been identified in individuals meeting clinical criteria for MCI [10] and have exhibited subtle relationships with cognition among older individuals without dementia or MCI symptoms [210]. In addition, pathogenic mutations in genes related to Aβ processing, including the amyloid precursor protein (*APP*) and presenilin-1 and -2 (*PSEN1*, *PSEN2*) genes, have been discovered in patients with the rare, autosomal dominant form of AD [13]. As a result, Aβ accumulation is increasingly proposed as a major antecedent ultimately leading to incident AD [2].

The fundamental biological influences on brain Aβ levels are not yet fully understood. The strongest known genetic risk factor for AD is presence of the *APOE* ε4 allele [14], and *in vitro* and murine studies have proposed plausible links between *APOE* ε4 and aberrant Aβ mechanisms [211]. However, *APOE* ε4 is neither necessary nor sufficient for development of AD pathology, suggesting that the biology underlying Aβ accumulation involves contributions from other genes and pathways, as well as the environment.

The ongoing search for genetic modulators of brain Aβ deposition in humans has been bolstered by advances in imaging methods for noninvasive detection of fibrillar Aβ *in vivo*. While existing genetic studies of brain Aβ have focused on candidate genes due to moderate sample sizes, the enhanced stability of recently-developed ^{18}F-labeled positron emission tomography (PET) imaging Aβ tracers such as florbetapir (also known as AV-45 or Amyvid) allows for more widespread evaluation to facilitate the acquisition of larger cohorts for analysis [212]. Importantly, florbetapir PET data has demonstrated strong relationships with pathologically-verified assessments of fibrillar Aβ burden [213, 214] and thus represents a novel and robust quantitative phenotype that can be assessed in samples with heightened power for discovery of genes influencing Aβ neuropathology.

We used quantitative florbetapir PET data from 555 participants enrolled in the ADNI cohort to perform the first GWAS of cortical Aβ burden in humans. We hypothesized that combining GWAS and genome-wide pathway analysis would confirm the association of *APOE* as well as identify other genetic modulators of brain Aβ deposition.

B. Methods

Study participants

This report utilized data from ADNI [162] (http://adni.loni.ucla.edu/), a multi-site longitudinal study that was launched in 2004 as a public-private partnership. The initial phase (ADNI-1) enrolled individuals aged 55-90 years who were recruited from over 50 sites across the United States and Canada and followed at 6- to 12-month intervals for 2-3 years. These individuals

included approximately 200 healthy controls (HC), 400 patients with late MCI (LMCI), and 200 patients clinically diagnosed with probable AD. Subsequent phases (ADNI-GO and ADNI-2) have extended follow-up for existing participants and have enrolled additional individuals, including those meeting criteria for early MCI (EMCI). All participants provided written informed consent and study protocols were approved by each site's institutional review board. Further information about ADNI, including full study protocols, complete inclusion and exclusion criteria, and data collection and availability can be found at http://www.adni-info.org/.

Florbetapir PET scans

PET imaging using the ^{18}F-labeled Aβ tracer florbetapir was performed for participants enrolled in ADNI-GO or ADNI-2. Participants were administered a bolus injection of approximately 370 MBq florbetapir intravenously. Fifty minutes later, a 20-minute continuous cranial PET scan was initiated. Images were reconstructed immediately following the scan using iterative algorithms, and repeat scans were acquired if motion artifact was detected. Preprocessing of the scans was performed as previously described [215]. Briefly, image frames were averaged, aligned to a standard space (AC-PC), resampled to a standard image and voxel size, smoothed to a uniform resolution, and normalized to an atlas-based bilateral and symmetric cerebellar reference region. This cerebellar reference region consisted largely but not exclusively of grey matter and was expected to exhibit nonspecific binding, ultimately resulting in standardized uptake value ratio (SUVR) images. These preprocessed scans were downloaded from the ADNI database (http://adni.loni.ucla.edu/) for 621 participants. For each scan, mean regional SUVR values were extracted for the frontal, parietal, temporal, limbic, and occipital lobes using the MarsBaR toolbox implemented in the Statistical Parametric Mapping 8 (SPM8) software

(http://www.fil.ion.ucl.ac.uk/spm/software/spm8/). The average SUVR for these 5 regions was then calculated to represent a global cortical measure of Aβ deposition to be used as a quantitative phenotype for GWAS. Overall, 19 participants (11 HC, 6 MCI, 2 AD) were excluded due to missing scan data or failed processing, leaving data for 602 individuals available for further analysis.

Genotyping and imputation

A blood draw for genomic DNA extraction was obtained at the screening or baseline visit for all study participants [170]. Genotyping on these samples was performed according to manufacturer's protocol (Illumina, Inc., San Diego, CA) using the Human610-Quad BeadChip (for subjects initially enrolled during ADNI-1) or the Human OmniExpress BeadChip (for subjects initially enrolled in ADNI-GO or ADNI-2). In addition, the two SNPs characterizing *APOE* ε2/ε3/ε4 status (rs429358 and rs7412) were genotyped separately as previously described [170] and merged with the array data sets. All genotype data underwent stringent quality control procedures using PLINK [94]. These steps included sample exclusion for call rate < 95% or failed identity or gender check, and SNP exclusion for call rate < 95%, Hardy-Weinberg equilibrium test $p < 1 \times 10^{-6}$, or MAF < 1% [170]. In addition, to limit possible effects of population stratification, multidimensional clustering analysis was used to select only participants with non-Hispanic Caucasian (CEU or TSI) ancestry based on HapMap3 reference populations. Overall, one individual was excluded due to a failed gender check and 42 individuals were excluded based on ancestry.

Next, haplotype patterns from the 1000 Genomes Project reference panel were used to impute genotypes for markers not directly assayed. Prior to imputation, the orientation of all genotyped markers in relation to the plus strand alignment of the reference panel genome (NCBI build 37 coordinates) was verified and monomorphic variants from the reference panel were excluded. Minimac [216] was used to impute samples within groups based on the genotyping platform employed (Illumina 610-Quad or OmniExpress). Following imputation, SNPs with $r^2 < 0.5$ between imputed and assayed genotypes were removed [217]. The remaining array SNPs demonstrated > 99.9% concordance between imputed and assayed genotypes.

The independently-imputed data sets were then merged to generate a common set of more than 10 million SNPs for the full ADNI sample. Following quality control (SNP call rate < 95%, Hardy-Weinberg $p < 1 \times 10^{-6}$) and frequency filtering (MAF < 5%), 6,108,668 SNPs were included in the GWAS. Of the 602 participants with Aβ PET data, 559 individuals were included in the resulting genetic data set. Among these individuals, four pairs exhibited significant relatedness (PLINK identity by descent PI_HAT > 0.5) and therefore one individual from each pair was randomly selected for exclusion (2 HC, 1 EMCI, 1 LMCI), leaving 555 participants for the final GWAS sample.

Statistical analysis

For the GWAS, linear regression was performed using PLINK to determine the association of each SNP to global cortical Aβ levels. An additive genetic model was specified and age, gender, and diagnosis (through a set of binary dummy variables indicating HC, EMCI, LMCI, or AD) were

applied as covariates. To account for multiple comparisons, we employed a conservative threshold for genome-wide significant association ($p < 5 \times 10^{-8}$) based on a Bonferroni correction of one million independent tests [218]. Haploview [180] was used to generate Manhattan and Q-Q plots and SNAP [219] and LocusZoom [220] were used to obtain regional association plots for selected loci. Post-hoc analyses, including hierarchical linear regression, effect size calculations, and exploratory correlation and interaction studies using Bonferroni corrections for multiple comparisons, were performed using IBM SPSS 20.0.

To extend the GWAS findings, we performed biological pathway enrichment analysis. We used GATES [221] to calculate a *p*-value for each gene accounting for its size, LD structure, and constituent SNP associations. MetaCore (GeneGo, Inc.) was used to identify pathways enriched with genes showing trend of association (GATES $p < 0.1$), applying a conservative threshold for pathway-level significance (FDR-corrected $p < 0.01$).

C. Results

This study analyzed data from 555 ADNI participants with non-Hispanic Caucasian ancestry. Participants were diagnosed at the time of the PET scan as HC, EMCI, LMCI, or AD. EMCI subjects met clinical criteria for amnestic MCI [9] but exhibited milder (between 1.0 and 1.5 standard deviations below age-associated norms) memory impairment (Table 6). In this sample, the EMCI group was younger than the other groups ($p < 0.001$). All groups displayed comparable levels of education but the LMCI group included fewer female participants compared to the HC group (Chi-square $p = 0.029$).

Table 6. Selected characteristics of ADNI participants at the time of PET scan. Data are number (%) or mean (SD). Abbreviations: CDR-SOB = Clinical Dementia Rating-Sum of Boxes; WMS-R = Wechsler Memory Scale-Revised.

	HC (n=179)	**EMCI (n=190)**	**LMCI (n=115)**	**AD (n=71)**
Age (years)	76.68 (6.25)	71.04 (7.41)	75.61 (8.14)	75.87 (8.15)
Women	87 (49%)	83 (44%)	41 (36%)	27 (38%)
Education (years)	16.27 (2.72)	15.89 (2.65)	16.11 (2.90)	16.04 (2.87)
***APOE* ε4 allele present**	41 (23%)	77 (41%)	49 (43%)	45 (64%)
CDR-SOB	0.07 (0.29)	1.22 (0.73)	1.73 (1.18)	5.63 (2.70)
Mini Mental Status Examination	29.07 (1.25)	28.39 (1.52)	27.74 (1.84)	21.68 (4.24)
Logical memory immediate recall (WMS-R)	14.94 (3.36)	10.93 (2.81)	8.74 (4.35)	4.20 (3.10)
Logical memory delayed recall (WMS-R)	14.08 (3.64)	8.87 (1.73)	6.13 (4.38)	1.67 (2.50)

The GWAS results did not indicate evidence of spurious inflation of association test statistics (λ = 1.00) due to population stratification or other confounding factors (Figure 9). Loci on two chromosomes exhibited genome-wide significant association ($p < 5 \times 10^{-8}$) to cortical Aβ levels (Figure 10). As expected, the peak association originated on chromosome 19 from rs429358 ($p = 5.45 \times 10^{-14}$), which is one of the two SNPs coding for the *APOE* ε4 allele [14]. While other SNPs within *APOE* and in the region of its adjacent genes *APOC1*, *TOMM40*, and *PVRL2* also displayed significant association to cortical Aβ levels in the primary GWAS model (Figure 10 and Figure 11A), their association signals disappeared ($p > 0.05$) when *APOE* ε4 status (absence = 0, presence = 1) was included as a covariate.

Multiple SNPs on chromosome 3 near *BCHE* (butyrylcholinesterase) also displayed genome-wide significant association to cortical Aβ load (Figure 10). The peak association signal at this locus originated from rs509208 ($p = 2.69 \times 10^{-8}$), which is approximately 450 kb upstream (5') of *BCHE* (Figure 11B). This association remained strong with the inclusion of *APOE* ε4 status as a covariate ($p = 1.94 \times 10^{-7}$).

Several additional loci exhibited suggestive association ($p < 5 \times 10^{-6}$) to cortical Aβ levels (Figure 10 and Table 7). These loci included SNPs within or near the cell adhesion genes of *ITGA6* (integrin, alpha 6; chromosome 2) and *ITGA1* (integrin, alpha 1; chromosome 5) as well as SNPs near the insulin signaling pathway gene *PIK3R1* (phosphoinositide-3-kinase, regulatory subunit 1; chromosome 5), among others. The association signals at these loci were not reduced after the inclusion of *APOE* ε4 allele status as a covariate (data not shown).

Figure 9. Q-Q plot of observed -log$_{10}$ *p*-values from the GWAS of cortical Aβ load versus those expected under the null hypothesis. The Q-Q plot exhibits no evidence of genomic inflation (PLINK-calculated λ = 1.00) or population stratification in the GWAS. An additive genetic model was used and age, gender, and diagnosis were applied as covariates. Analyses were restricted to subjects with non-Hispanic Caucasian (CEU or TSI) ancestry as determined by genetic clustering. Observed -log$_{10}$ *p*-values > 8 are represented along the top of the plot as red triangles, while all other values are represented as red dots.

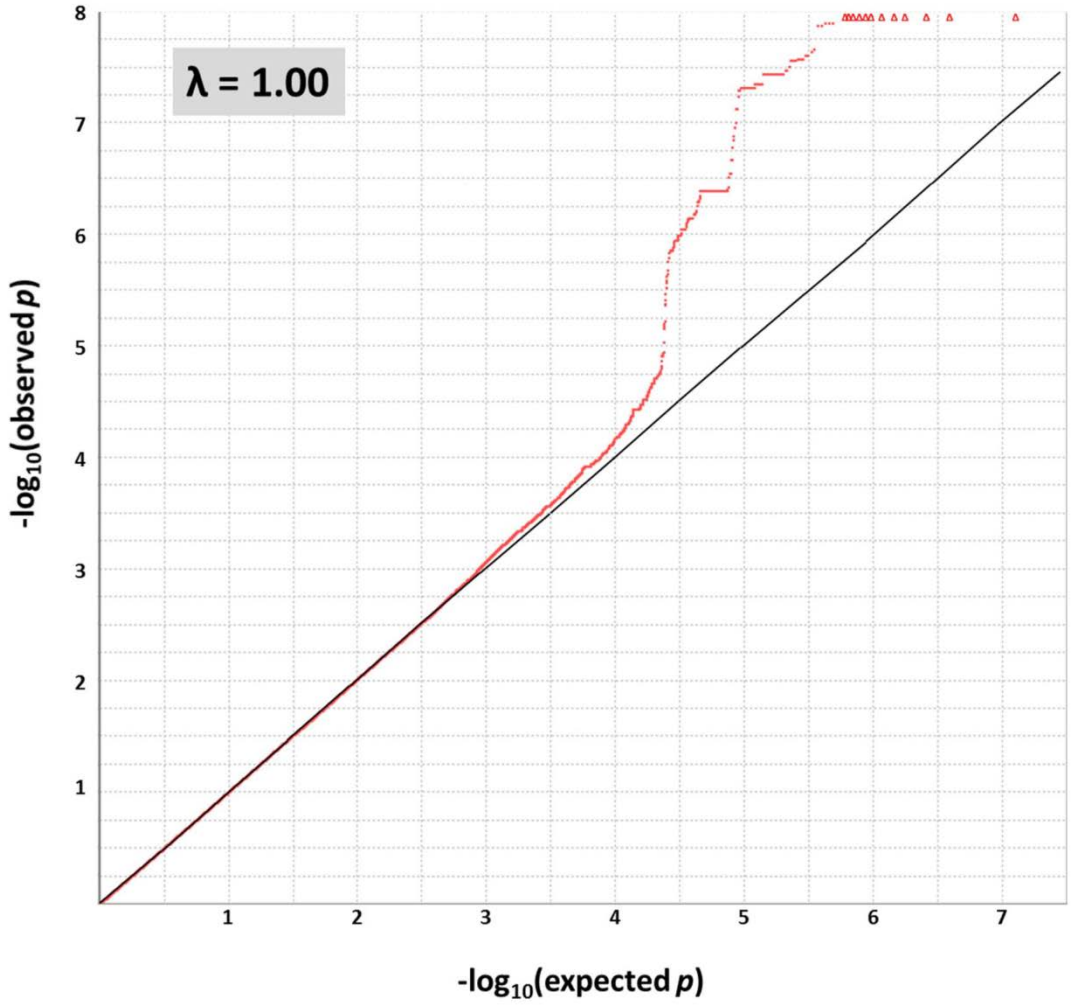


Figure 10. Manhattan plot of observed -log$_{10}$ *p*-values from the GWAS of cortical Aβ load. More than six million SNPs were tested for association to global cortical Aβ burden under an additive genetic model and applying age, gender, and diagnosis as covariates. Genome-wide significant associations (exceeding the threshold represented by the red line) were identified on chromosome 19 within *APOE* and its neighboring genes and on chromosome 3 at the *BCHE* locus. Suggestive associations (exceeding the threshold represented by the blue line) were identified on five additional chromosomes. Annotations are provided for genome-wide significant associations and for the top three suggestive associations.

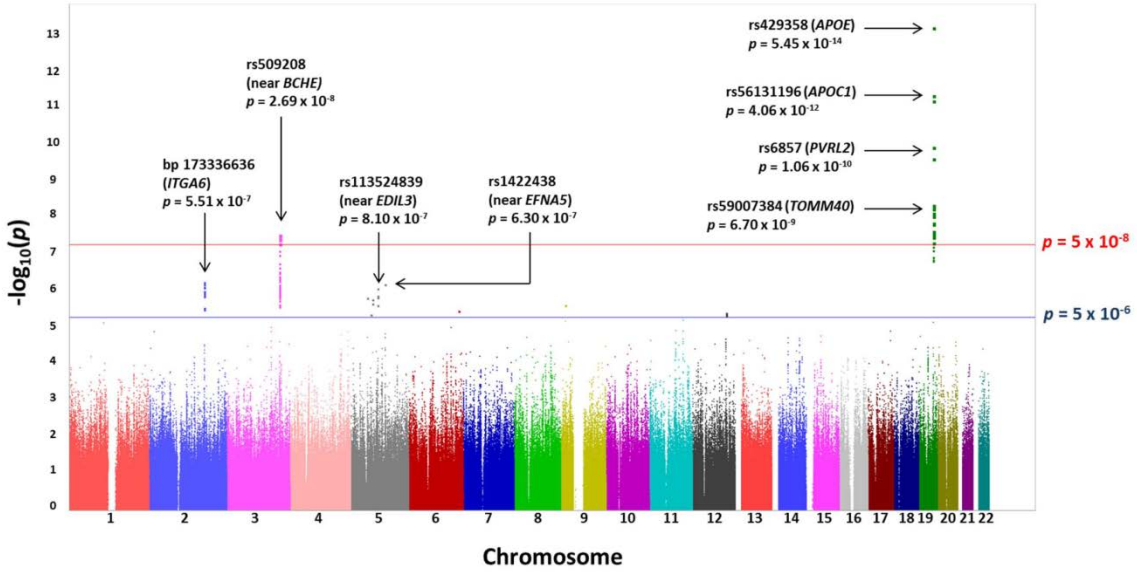


Figure 11. Regional association plots for the loci exhibiting genome-wide significant association to cortical Aβ burden. Magnified association plots are displayed for the regions around A) rs429358 within *APOE* and B) rs509208 at the *BCHE* locus. SNPs are plotted based on their GWAS -$\log_{10}$ *p*-values (left vertical axis) and their genomic position (NCBI build 36). Genes in these regions are labeled with arrows denoting their 5'-to-3' orientation, and the red color scale of r^2 values is used to label SNPs based on their degree of LD with the annotated peak SNP.

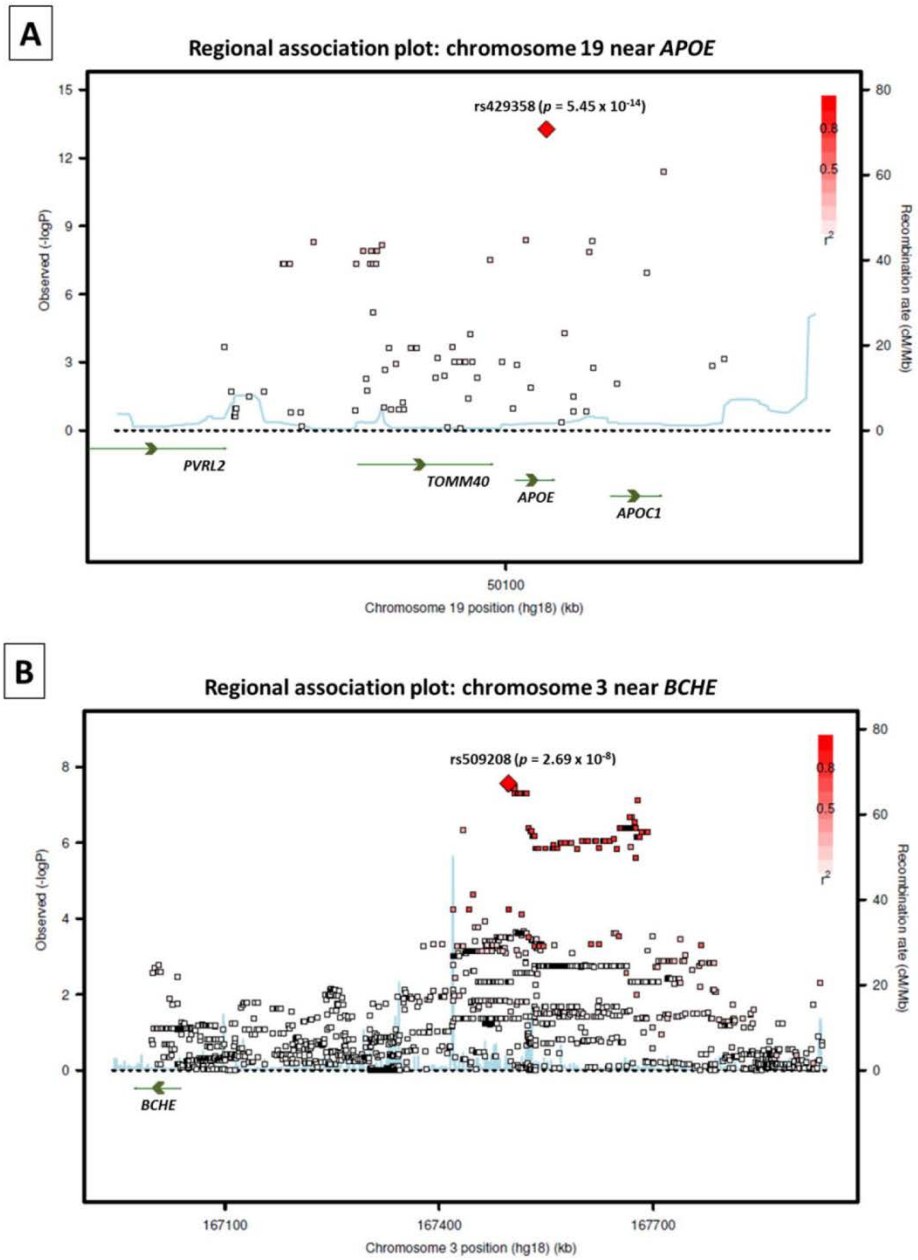


Table 7. Peak association signals ($p < 5 \times 10^{-6}$) from unique genes or intergenic loci in the GWAS of global cortical Aβ load.

Chr	SNP Identifier	Gene	Minor Allele (MAF)	*P*-value
19	rs429358	*APOE*	C (0.28)	5.45×10^{-14}
19	rs56131196	*APOC1*	A (0.29)	4.06×10^{-12}
19	rs6857	*PVRL2*	T (0.28)	1.06×10^{-10}
19	rs59007384	*TOMM40*	T (0.32)	6.70×10^{-9}
3	rs509208	5' of *BCHE*	G (0.16)	2.69×10^{-8}
2	Position 173336636 (no dbSNP ID)	*ITGA6*	A (0.43)	5.51×10^{-7}
5	rs1422438	5' of *EFNA5*	G (0.30)	6.30×10^{-7}
5	rs113524839	5' of *EDIL3*	T (0.11)	8.10×10^{-7}
5	rs7702276	5' of *ITGA1*	T (0.36)	1.47×10^{-6}
5	rs24449894	5' of *PIK3R1*	A (0.17)	1.62×10^{-6}
9	rs7039300	3' of *NFIB*	G (0.23)	2.37×10^{-6}
6	rs9384488	5' of *ARID1B*	A (0.35)	3.39×10^{-6}
12	rs10219670	Between *NUAK1* and *C12orf75*	C (0.42)	3.89×10^{-6}

Following the GWAS, we performed post-hoc analyses to further assess the impact of the *APOE* and *BCHE* loci on Aβ burden. While both the *APOE* ε4 allele and the minor allele (G) of rs509208 conferred increases in cortical Aβ levels (Figure 12A and 12B), there was no evidence of epistasis modeled as an interaction between these factors ($p = 0.871$). Instead, these factors appeared to exert independent and additive effects on Aβ burden (Figure 12C), with comparable effect size associated with presence of at least one copy of the minor allele at rs509208 whether *APOE* ε4 allele status was included as a covariate (Cohen's $d = 0.50$) or not (Cohen's $d = 0.52$). Exploratory analyses did not reveal significant interactions ($p > 0.05$) of the *APOE* or *BCHE* risk loci with age, diagnosis, education, or gender.

We next performed hierarchical linear regression to assess the variance in Aβ levels uniquely explained by these genetic factors (ΔR^2). Age, gender, and diagnosis were entered as the first block in the model, and collectively accounted for 2.7% of the variance in cortical Aβ levels in this sample. *APOE* ε4 allele status (+/-) and *BCHE* rs509208 allele status (+/-) were included in the second block for stepwise entry into the model. *APOE* ε4 was found to explain an additional 10.7% of the variance, while *BCHE* rs509208 accounted for 4.3% of variance over and above that explained by the previously entered variables.

Several biological pathways exhibited enrichment of association to Aβ deposition (Table 8). Among these included several pathways related to signaling through NOTCH, opioid receptors, and the epidermal growth factor receptor (EGFR). Pathways related to mitogen-activated protein kinase signaling (MAPK; also known as ERK), cell adhesion, and activation of estrogen receptors also displayed enrichment of association.

Figure 12. *APOE* ε4 and rs509208 (*BCHE*) appear to exhibit independent, additive effects on cortical Aβ levels. Mean cortical Aβ levels (adjusted for age, gender, and diagnosis) ± standard errors are displayed based on A) the number of *APOE* ε4 allele copies and B) rs509208 genotype. Presence of at least one copy of the ε4 allele was significantly associated with increased Aβ burden (Cohen's *d* = 0.71), as was presence of at least one copy of the minor allele (G) of rs509208 (Cohen's *d* = 0.52). These loci appeared to exert additive effects on Aβ levels (panel C): subjects having both risk factors exhibited significantly greater Aβ burden than subjects having either factor in isolation, and no significant epistasis modeled as an interaction was identified.

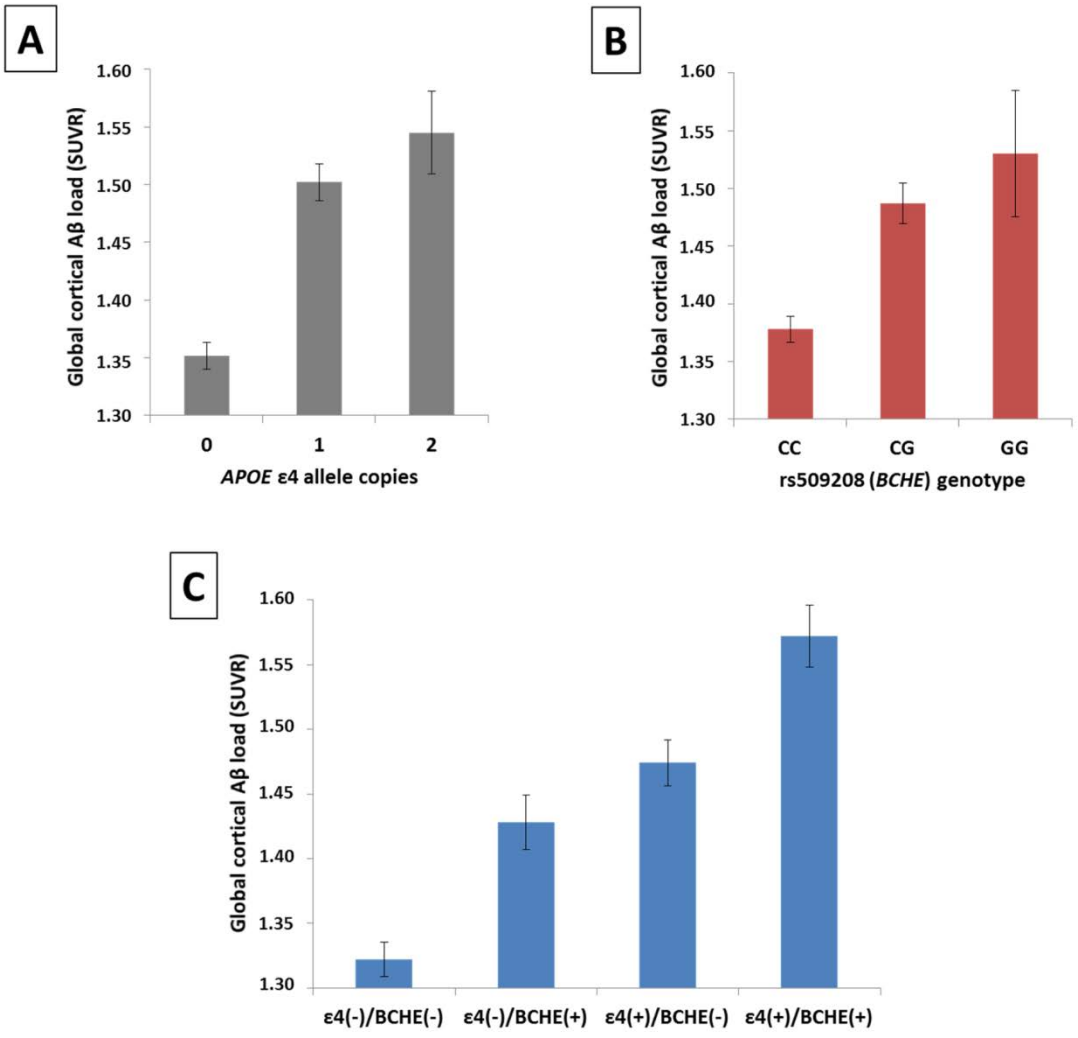


Table 8. Pathways showing enrichment of association (FDR $p < 0.01$) to cortical Aβ load.

Pathway	Genes in pathway	Nominal *p*
Development: Delta- and kappa-type opioid receptors signaling via beta-arrestin	23	1.91×10^{-5}
Development: Dopamine D2 receptor transactivation of EGFR	24	2.60×10^{-5}
Development: Notch Signaling Pathway	43	3.41×10^{-5}
Development: NOTCH1-mediated pathway for NF-KB activity modulation	34	3.81×10^{-5}
Apoptosis and survival: Role of CDK5 in neuronal death and survival	34	3.81×10^{-5}
G-protein signaling: G-Protein alpha-q signaling cascades	34	3.81×10^{-5}
Development: Ligand-independent activation of ESR1 and ESR2	45	4.99×10^{-5}
Development: NOTCH-induced EMT	19	6.11×10^{-5}
Development: Mu-type opioid receptor regulation of proliferation	28	7.77×10^{-5}
Development: Gastrin in differentiation of the gastric mucosa	38	8.96×10^{-5}
Development: ERBB-family signaling	39	1.09×10^{-4}
Development: EGFR signaling pathway	63	1.50×10^{-4}
Cell adhesion: ECM remodeling	52	1.63×10^{-4}
Development: ACM2 and ACM4 activation of ERK	43	2.24×10^{-4}
Development: EGFR signaling via small GTPases	33	2.37×10^{-4}

D. Discussion

Using florbetapir PET and GWAS, this study confirmed the association of *APOE* and identified a novel and independent association of *BCHE* to cerebral Aβ deposition. Together, the risk variants at these loci explained 15% of the variation in cortical Aβ levels, a substantial effect for two genes in a study of this size of complex disease. Additional loci, including both new genes and others previously studied in relation to AD, displayed suggestive association to Aβ burden and provide further targets for follow-up. Finally, numerous biological pathways exhibited enrichment of association to cortical Aβ load, including several targets under active investigation for drug development.

Florbetapir PET allows for noninvasive detection of brain Aβ plaques, a hallmark pathologic feature of AD [213, 214]. It also serves as a quantitative endophenotype that can provide increased statistical power for discovery using GWAS compared to case-control designs [12]. Although the heritability of Aβ deposition, a dynamic process captured by PET at one time point, is unknown and not a direct proxy for AD heritability, implicated markers using this approach may be more closely related to underlying processes impacting disease risk and progression compared to markers discovered using case-control designs [11].

For late-onset AD, the largest known genetic risk factor is the *APOE* ε4 allele. Extensive prior studies have suggested a number of mechanistic roles for *APOE* ε4 on Aβ burden, including hindering clearance of soluble Aβ from the brain [222], favoring Aβ aggregation into fibrils [223], and promoting neurodegeneration by directing toxic Aβ oligomers to synapses [224]. Among the genes neighboring *APOE*, we found no significant associations with Aβ burden after

including *APOE* ε4 status as a covariate. While this suggests that these genes did not exhibit independent effects on Aβ deposition, the extensive LD structure around *APOE* makes this difficult to definitively determine.

BCHE has previously been proposed as an AD risk gene [225]. Butyrylcholinesterase is known to be enriched within Aβ plaques in post-mortem human AD brains [226], and its increased presence has been suggested as a critical factor in the formation of the neuritic plaques of dementia [227]. The most commonly studied *BCHE* SNP is the K-variant (rs1803274), which is approximately 500 kb downstream of and not in LD with rs509208. The *BCHE* K-variant has demonstrated synergistic effects with *APOE* ε4 on incidence of pathologically-confirmed late-onset AD [225] and on risk of progression from MCI to AD [228]. Nevertheless, the present study is the first to implicate genetic variation at the *BCHE* locus in brain Aβ burden in humans and represents the largest reported effect for this gene on an AD-related phenotype.

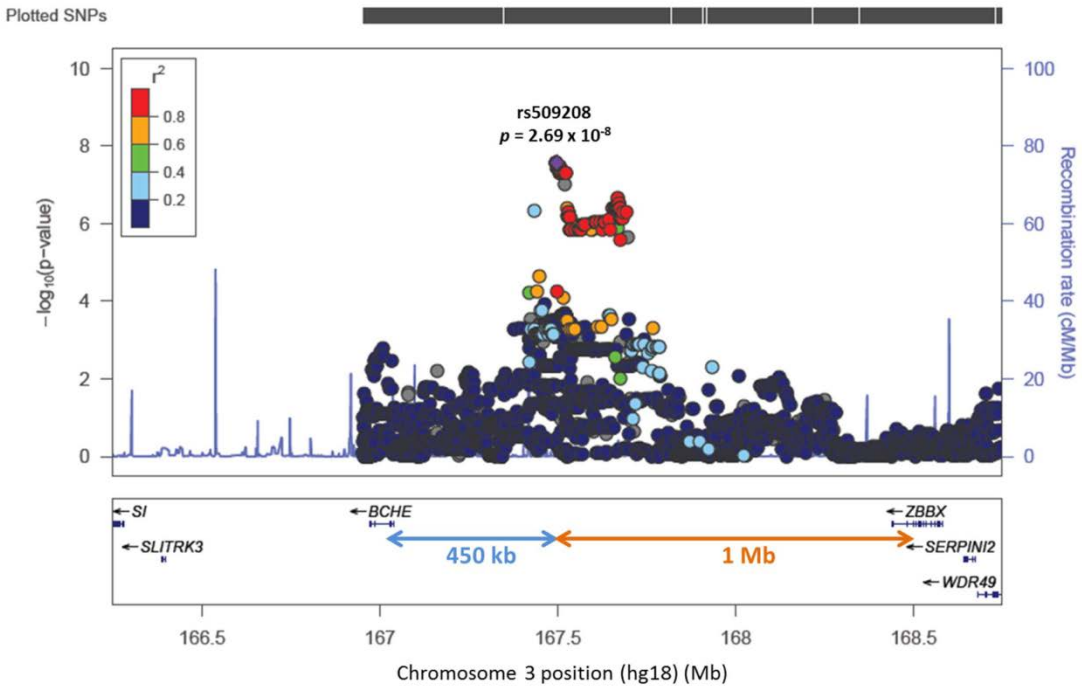
There are several mechanisms which might explain the effect of *BCHE* on Aβ plaque burden. Genetic variation at *BCHE* has been associated with increased cortical butyrylcholinesterase activity in autopsy tissue from elderly individuals with dementia [229]. Increased enzyme activity, leading to decreased acetylcholine levels, may disrupt synaptic functioning, glial cell activation, and myelin maintenance to favor Aβ plaque formation and neurodegeneration [230]. Indeed, cholinesterase inhibitors are presently first-line symptomatic therapies for AD [5], and drugs such as rivastigmine which inhibit butyrylcholinesterase are suggested to have potential disease-modifying effects in certain individuals compared to exclusive acetylcholinesterase inhibitors [231]. Alternatively, through non-enzymatic functions butyrylcholinesterase may promote Aβ fibrillogenesis from soluble precursors [232] or may interact with Aβ and *APOE* to

alter the cerebrospinal fluid environment [233, 234] and to render developing plaques more resistant to clearance [235]. Interestingly, butyrylcholinesterase activity has been associated with insulin resistance [236], and rs509208 in particular has been associated in a separate GWAS with diabetes-related traits [237], suggesting a broader role of the *BCHE* locus in disorders characterized by amyloidogenic protein accumulation.

Although rs509208 is approximately 450 kb upstream (5') of *BCHE* and not within conventional gene boundaries, the next closest genes (*ZBBX*, *SERPINI2*) are nearly 1 Mb in the opposite direction (Figure 13). Of note, SNPs as far as 800 kb upstream of *BCHE* have previously demonstrated genome-wide significant association with serum butyrylcholinesterase activity in a population sample of nearly 9000 individuals from several Australian twin and family studies [238]. These SNPs included a peak signal 250 kb from the gene (rs2034445) and other non-independent signals from SNPs in high LD with rs509208 (e.g., rs6443374 and rs13314077), suggesting that variants upstream of the gene may exert regulatory effects on *BCHE* expression with consequences to the activity of its encoded enzyme. Converging evidence in genomics indicates that this kind of regulation is quite common and may involve mechanisms influencing chromatin structure, transcription factor binding, and splicing component recognition sequences, among others [89]. Molecular characterization in brain tissue and cell cultures will be important to determine the complex functional architecture of the *BCHE* locus and the genes and other DNA elements surrounding it.

Biological pathways provide additional vehicles for characterizing complex genetic architectures, since variants of modest individual effect can collectively influence susceptibility through action within shared mechanisms [31]. We observed enrichment of association to Aβ deposition

Figure 13. Regional association plot (wide view) around rs509208 at the *BCHE* locus. SNPs within *BCHE* and the region spanning approximately 1.5 Mb upstream of *BCHE* are plotted based on their GWAS -$\log_{10}$ *p*-values (left vertical axis), NCBI build 36 genomic position (horizontal axis), and recombination rates calculated from 1000 Genomes Project reference data (right vertical axis). Genes are labeled with arrows denoting their 5'-to-3' orientation. The color scale of r^2 values is used to label SNPs based on their degree of linkage disequilibrium with rs509208. As displayed, rs509208 is 450 kb upstream of *BCHE* (blue arrows) and is approximately 1 Mb away from the next closest genes (orange arrows).



within multiple pathways related to EGFR signaling, which has been proposed as a target for treating amyloid-β-induced memory loss [239], as well as NOTCH signaling, which contributes to neuronal plasticity and interacts with several Aβ-generating mechanisms [240, 241]. Dysfunction within opioid receptor signaling may also promote the generation of Aβ through up-regulating enzymes that favor Aβ formation [242] and through indirect effects on cyclic AMP response-element binding protein (CREB), which is a key molecular switch mediating long-term memory formation [143]. Genetic variation in cell adhesion genes has also been previously implicated in AD risk, AD-related memory impairment, and Aβ generation in animal models [35, 59, 189, 243], but have not been previously linked to Aβ load in humans. Together, these results provide additional mechanistic insights into Aβ deposition and further validate the use of pathway-based approaches to detect robust effects in GWAS data that are otherwise concealed beneath the stringent thresholds for genome-wide SNP-level significance.

The current study has several limitations. Although this represents the largest genetic study of Aβ PET, the sample size has limited power for a GWAS. With a larger sample, the suggestive loci we highlighted might have achieved genome-wide significance. Among these included *ITGA6* which encodes a component of a receptor complex proposed to mediate the pro-inflammatory interaction of microglia with Aβ fibrils [244], *PIK3R1* which may contribute to disruptions in insulin signaling in the AD brain, [245] and other genes with potential relationships to AD pathogenesis. Similarly, given the novelty of the present Aβ PET GWAS data set, a comparable replication sample is not yet available.

In addition, the ADNI cohort represents a sample typical of a clinical trial for MCI/AD and the analyses here were restricted to non-Hispanic Caucasians. The extent to which the present

findings can be generalized to other populations and to the community setting of MCI/AD individuals remains to be determined. Future multi-center and international collaborations are expected to yield larger samples with greater power for analyses of potential interactions of genetic risk factors with each other and with clinical variables such as gender, ethnicity, family history, age of onset, and rate of disease progression, which could not be appropriately addressed with presently available data. Florbetapir also does not bind soluble forms of Aβ [246] which may exhibit dynamic relationships with deposited, fibrillar Aβ to drive neurodegeneration in AD [208]. Finally, while this study employed imputed (probabilistically-predicted) genotype data to provide deep coverage of the genome, higher-density genotyping arrays and sequencing will eventually provide direct assays for a similar number of variants.

Despite these limitations, the present findings point to several intriguing extensions for follow-up. First, GWAS of longitudinal change in florbetapir PET Aβ burden is likely to elucidate additional genes modulating the rate of progression of AD neuropathology. Complementary analytical strategies, including network- and epistasis-based approaches, may also reveal functional influences on Aβ deposition that are not easily observed through GWAS and pathway analysis. In addition, whole genome sequencing will dramatically enhance the granularity of coverage for GWAS-implicated loci and could be particularly valuable for discovering additional novel loci, rare variants, copy number variants, and DNA regulatory elements. Finally, the approval for widespread use of both florbetapir PET imaging and butyrylcholinesterase inhibitors creates a unique opportunity to prospectively assess the effects of these drugs on Aβ deposition over time, particularly among individuals at early stages in the AD spectrum where clinical efficacy would likely be most valuable.

This study highlights the power of pairing targeted molecular imaging with genome-wide analytical strategies to elucidate mechanisms underlying AD pathophysiology. The associations of the *BCHE* locus and the NOTCH and EGFR pathways to Aβ deposition merit further investigation and may have significant implications for risk stratification, biomarker interpretation, and therapeutic development.

V. Conclusions and future directions

In this work, we critically reviewed and synthesized strategies for pathway analysis of genomic data and applied pathway analysis as a complementary approach to GWAS of episodic memory performance and cerebral Aβ deposition in AD. Through this framework, we confirmed the association of *APOE* (apolipoprotein E) to these key AD endophenotypes and discovered additional genes and pathways modulating memory functioning and Aβ pathology in AD. Our findings further validate the promise of pathway-based analyses for disorders with complex genetic architectures and nominate or highlight several intriguing biological mechanisms for further study in AD.

Interestingly, we observed several genes and pathways which appear to bridge the mechanisms underlying memory and cognitive functioning on one hand and Aβ burden and other measures of neuropathology on the other. For example, *BCHE* has long been known to impact learning and memory through regulating levels of the neurotransmitter acetylcholine [226], and through GWAS in this work, has now been demonstrated to impact Aβ burden as well. Pathways related to cell adhesion, inflammation, and signaling through NOTCH, EGFR, and MAPK also showed association to both memory performance and Aβ deposition in our analyses. That these pathways displayed relationships to such distinct AD endophenotypes emphasizes that they are likely to play major roles in disease pathogenesis and as such, warrant further study in independent data sets and using alternative analytical strategies.

It is also noteworthy that the GWAS results from our analyses displayed quite different association profiles. For example, the GWAS of memory performance identified significant SNP-

level signal only in the *APOE* region, while the GWAS of Aβ deposition, despite including a smaller sample size, detected a novel genome-wide significant association of *BCHE* [247]. Nevertheless, pathway analysis successfully revealed additional associations in both cases (i.e., whether there was robust SNP-level signal and where there was not). These findings provide an empirical demonstration of the power of pathway analysis to complement GWAS by identifying broader functional trends that may not be obvious from top-line results alone.

Our results also suggest that future studies, as well as diagnosis and treatment strategies, may need to evolve to reflect a complex AD genetic architecture involving multiple pathways. One possibility is that integration of multiple clinical biomarkers – such as genotype, blood and CSF analyte, brain imaging, cognitive assessment, and medical history data – might be appropriate to detect replicable effects of key pathways. In addition, from a clinical standpoint, the functions of many disease pathways may be dynamic to disease stage. For example, high blood levels of a particular cytokine might have different implications for risk stratification depending on the genetic profiles of key pathways within the context of brain structure and other clinical measures. Similarly, therapeutic and preventative strategies for neurodegenerative disease may benefit from drug combinations based on the cocktail approaches used for HIV infection and some cancers. It is possible that efficacy, and therefore the choice of particular drugs to include in the cocktail, may depend on an individual's profile of biomarkers and key genetic variants – some of which may be protective and others deleterious – in targeted pathways. The development of advanced statistical models for analysis of large, multimodal data sets will help to explore these potentially new paradigms.

Notably, extensions of this work using next-generation sequencing may be uniquely powered to discover other associations to memory and Aβ pathology, including novel SNPs, rare, structural, and epigenetic variants, and dynamic molecular changes in the transcriptome [43]. Similar genetic studies of memory functioning and Aβ deposition in population-based samples may provide further insight into the delineations between normal cognitive aging and pathological declines in the AD spectrum. In addition, measures of longitudinal change in these endophenotypes may provide enhanced power for analysis and may also elucidate unique mechanistic influences. Further, although it was beyond the scope of this study, characterization of gene-environment interactions is an important under-investigated avenue that will likely guide the optimized application of drug and lifestyle modifiers to combat AD onset and progression.

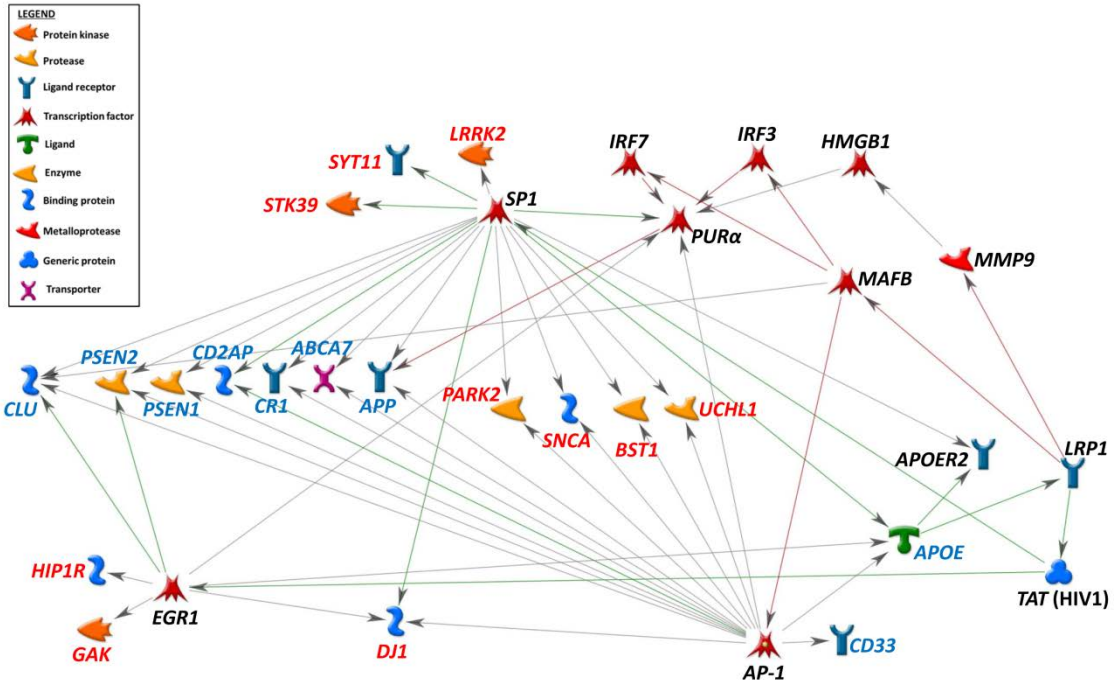
Pathway- and network-based approaches also have the potential to reveal shared mechanisms among AD and other neurodegenerative disorders. As an initial demonstration of this potential, we performed a preliminary network analysis to identify additional functional relationships between top AD- and PD-associated genes. Due to the numerous pathways implicated in AD and PD and the pleiotropic effects of many key disease-associated genes [248, 249], we hypothesized that regulatory relationships among these genes might impact multiple pathways and explored this hypothesis through transcription factor network analysis using the MetaCore software (GeneGo, Inc.). This approach incorporates knowledge from published literature to relate an input list of genes to known transcription factors and proximal targets such as ligand-receptor interactions. As input, we used the top 10 genes from the AlzGene (*APOE*, *BIN1*, *CLU*, *ABCA7*, *CR1*, *PICALM*, *MS4A6A*, *CD33*, *MS4A4E*, *CD2AP*) [127] and PDGene (*MAPT*, *SNCA*, *GBA*, *LRRK2*, *PM20D1*, *GAK*, *MCCC1*, *STK39*, *BST1*, *GPNMB*) [128] databases in addition to a small

number of genes (*APP*, *PSEN1*, *PSEN2*, *DJ1*, *HIP1R*, *PARK2*, *SYT11*, *UCHL*) implicated in both Mendelian and sporadic forms of AD or PD.

A network was identified which displays relationships among 31 factors, including 19 of the 28 input genes (Figure 14). The probability of the software algorithm generating a network with this level of interconnectedness by random selection of input genes was exceedingly small ($p = 1.14 \times 10^{-54}$). Strikingly, numerous genes in the network exhibit co-regulation by the SP1 (specificity protein 1) and AP-1 (activating protein 1) transcription factors. SP1 has been previously noted to regulate the expression of multiple AD-related genes, including a collection of memory-related genes in our earlier analyses [35, 205]. Elevated levels of SP1 have been identified in AD human brains and mouse models [250, 251] and may be induced by inflammation and oxidative stress [204, 250]. The AP-1 transcription factor is composed of heterodimers of several proteins, including those encoded by the *FOS* and *JUN* proto-oncogenes [252]. AP-1 is an important regulator of dopaminergic signaling pathways [253, 254] as well as numerous genes related to autophagy and lysosomal function [255]. Interestingly, animal models indicate that inhibition of SP1 may be neuroprotective in AD [251] and inhibition of AP-1 may be neuroprotective in PD [256]. The connections among SP1, AP-1, and AD- and PD-associated genes suggest that coordinated modulation of these transcription factors may be a viable strategy for combating neurodegeneration and merits further study.

This transcriptional network also includes several additional genes of interest which were not in the initial input list. For example, *EGR1* (early growth response 1) encodes a zinc-finger transcription factor that is important for synaptic plasticity [257] and cognitive performance [258] and whose up-regulation in AD brains may promote phosphorylation of tau [259]. The

Figure 14. Regulatory network centered on the SP1 and AP-1 transcription factors is enriched with top AD and PD genes. Meta-analytic genetic association data from public databases and supplementary manual curation was used to generate a list of 13 AD genes and 15 PD genes. Network analysis was performed using the MetaCore software to relate these input genes to known transcription factors and proximal targets based on published findings. A highly interconnected network including 9 AD genes (labeled in blue), 10 PD genes (labeled in red), and 13 additional genes (labeled in black) was identified. Many of the input AD and PD genes exhibit co-regulation by the SP1 and AP-1 transcription factors. Other genes of interest were also related to input AD and PD genes and represent a variety of candidate pathways in neurodegeneration.



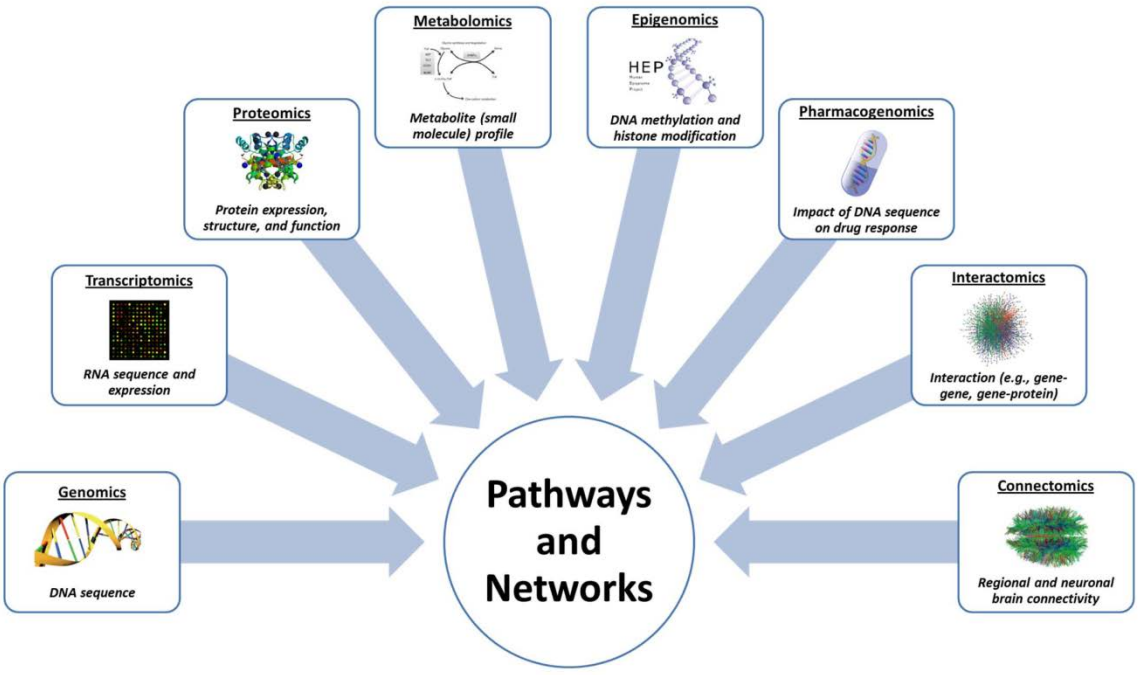
transcription factor encoded by *HMGB1* (high-mobility group protein 1) can also directly bind aggregates of the PD protein α-synuclein [260], regulate phagocytosis of Aβ [261, 262], and promote inflammation when secreted by activated microglia or necrotic neurons [263, 264]. Interactions between HIV-1 *TAT* (transactivator of transcription) and genes involved in AD and PD may be involved in HIV-associated cognitive impairment and Aβ pathology [265, 266]. Other genes of interest in this regulatory network include *MMP9* (matrix metalloproteinase 9) which is involved in synaptic plasticity and Aβ degradation [267], *IRF3* and *IRF7* (interferon regulatory factors 3 and 7) which regulate interferon-mediated inflammation and immune responses [268-271], and *LRP1* (low density lipoprotein receptor-related protein 1) which may affect several neurodegeneration pathways including lipid metabolism, Aβ endocytosis, and inflammation [272-275].

It should be noted that this type of analysis is preliminary and is not comprehensive or unbiased. Complementary strategies, including the use of alternative criteria for selection or statistical weighting of input genes as well as other schema for defining network connections, might highlight different relationships. Nevertheless, this regulatory network generates hypotheses for further investigation and reflects, at the transcriptional level, many of the same pathways implicated by genetic studies of AD and PD. More broadly, these findings argue for a better understanding of altered transcriptional regulation patterns through whole genome expression arrays and whole transcriptome sequencing (RNA-seq). These data modalities, particularly if viewed through a pathway- and network-based lens, would likely augment GWAS findings in AD, PD, and other neurodegenerative diseases by providing functional information to connect genetic variation with biochemical outcomes.

As an extension of this concept and the insights from the work described here, we propose that pathways and networks can serve as vehicles for integrating findings from diverse study modalities in AD and other complex disorders. There are many active strategies for large scale omics analysis (Figure 15), and findings that converge across these multiple study designs can provide confirmatory evidence that is crucial for efficient clinical translation. Isolated genes and molecules can be challenging to evaluate for convergence since they may not be represented in all data modalities or experimental model systems. In contrast, pathways and networks can incorporate data from multiple biological levels (e.g., genes, transcripts, proteins, and metabolites, among others) and may be more likely to be evolutionarily conserved [276]. For example, recent pathway-based studies integrating GWAS and gene expression data have demonstrated enhanced power, reproducibility, and connections of top findings to hypothesized disease processes [61, 91, 277].

The utility of pathway-based studies for AD and other complex disorders will continue to increase as present limitations in extant statistical approaches are addressed, including how to incorporate associations from intergenic regions and from genes without known functions. Nevertheless, it is clear that a pathway-based framework has significant strengths for AD genetic studies, including enhanced statistical power and the underlying emphasis that discoveries of strongly-associated genes represent a foundation to study their larger functional environment, since other components in that environment may yield superior targets for biomarker and drug development [41, 42, 141]. These advantages will be vital in harnessing the wealth of existing data on AD to develop an integrated understanding of its mechanisms and formulate optimal strategies for clinical translation.

Figure 15. Biological pathways and networks: a hub for convergent omics. Numerous large scale omics approaches are being used to study complex neurodegenerative diseases and endophenotypes in human tissue and animal and other model systems. Unlike individual genes and other isolated molecules, which may not be present in all model systems and may have differential sensitivity for detection with various study designs, pathways and networks are well-conserved and can be evaluated for convergence across diverse methodological approaches. Integration of findings to identify pathways and networks with consistent relationships to disease is likely to enhance the development of diagnostic biomarkers and treatment and prevention strategies.



References

1. Nussbaum, R.L. and C.E. Ellis, *Alzheimer's Disease and Parkinson's Disease.* New England Journal of Medicine, 2003. **348**(14): p. 1356-1364.

2. Sperling, R.A., et al., *Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease.* Alzheimers Dement, 2011. **7**(3): p. 280-92.

3. Jack, C.R., et al., *Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers.* The Lancet Neurology, 2013. **12**(2): p. 207-216.

4. Hurd, M.D., et al., *Monetary costs of dementia in the United States.* N Engl J Med, 2013. **368**(14): p. 1326-34.

5. Farlow, M.R., M.L. Miller, and V. Pejovic, *Treatment options in Alzheimer's disease: maximizing benefit, managing expectations.* Dement Geriatr Cogn Disord, 2008. **25**(5): p. 408-22.

6. Krstic, D. and I. Knuesel, *Deciphering the mechanism underlying late-onset Alzheimer disease.* Nat Rev Neurol, 2013. **9**(1): p. 25-34.

7. Mayeux, R., *Clinical practice. Early Alzheimer's disease.* N Engl J Med, 2010. **362**(23): p. 2194-201.

8. Saykin, A.J., et al., *Older adults with cognitive complaints show brain atrophy similar to that of amnestic MCI.* Neurology, 2006. **67**(5): p. 834-42.

9. Petersen, R.C., *Clinical practice. Mild cognitive impairment.* N Engl J Med, 2011. **364**(23): p. 2227-34.

10. Petersen, R.C., et al., *Neuropathologic features of amnestic mild cognitive impairment.* Arch Neurol, 2006. **63**(5): p. 665-72.

11. Kendler, K.S. and M.C. Neale, *Endophenotype: a conceptual analysis.* Mol Psychiatry, 2010. **15**(8): p. 789-97.

12. Potkin, S.G., et al., *Genome-wide strategies for discovering genetic influences on cognition and cognitive disorders: methodological considerations.* Cogn Neuropsychiatry, 2009. **14**(4-5): p. 391-418.

13. Sleegers, K., et al., *The pursuit of susceptibility genes for Alzheimer's disease: progress and prospects.* Trends Genet, 2010. **26**(2): p. 84-93.

14. Corder, E.H., et al., *Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families.* Science, 1993. **261**(5123): p. 921-3.

15. Barral, S., et al., *Genotype patterns at PICALM, CR1, BIN1, CLU, and APOE genes are associated with episodic memory.* Neurology, 2012. **78**(19): p. 1464-71.

16. Chibnik, L.B., et al., *CR1 is associated with amyloid plaque burden and age-related cognitive decline.* Ann Neurol, 2011. **69**(3): p. 560-9.

17. Hu, X., et al., *Meta-analysis for genome-wide association study identifies multiple variants at the BIN1 locus associated with late-onset Alzheimer's disease.* PLoS One, 2011. **6**(2): p. e16616.

18. Jun, G., et al., *Meta-analysis Confirms CR1, CLU, and PICALM as Alzheimer Disease Risk Loci and Reveals Interactions With APOE Genotypes.* Arch Neurol, 2010. **67**(12): p. 1473-1484.

19. Keenan, B.T., et al., *A coding variant in CR1 interacts with APOE-ε4 to influence cognitive decline.* Human Molecular Genetics, 2012.

20. Lambert, J.-C., et al., *Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease.* Nat Genet, 2009. **41**(10): p. 1094-1099.

21. Guerreiro, R., et al., *TREM2 variants in Alzheimer's disease.* N Engl J Med, 2013. **368**(2): p. 117-27.

22. Jonsson, T., et al., *Variant of TREM2 associated with the risk of Alzheimer's disease.* N Engl J Med, 2013. **368**(2): p. 107-16.

23. Hollingworth, P., et al., *Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease.* Nat Genet, 2011. **advance online publication**.

24. Bis, J.C., et al., *Common variants at 12q14 and 12q24 are associated with hippocampal volume.* Nature Genetics, 2012. **44**(5): p. 545-551.

25. Pottier, C., et al., *High frequency of potentially pathogenic SORL1 mutations in autosomal dominant early-onset Alzheimer disease.* Mol Psychiatry, 2012. **17**(9): p. 875-9.

26. Naj, A.C., et al., *Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease.* Nat Genet, 2011.

27. Gatz, M., et al., *Role of genes and environments for explaining Alzheimer disease.* Arch Gen Psychiatry, 2006. **63**(2): p. 168-74.

28. Kamboh, M.I., et al., *Genome-wide association study of Alzheimer's disease.* Transl Psychiatry, 2012. **2**: p. e117.

29. Simon-Sanchez, J. and A. Singleton, *Genome-wide association studies in neurological disorders.* Lancet Neurol, 2008. **7**(11): p. 1067-72.

30. Cantor, R.M., K. Lange, and J.S. Sinsheimer, *Prioritizing GWAS results: A review of statistical methods and recommendations for their application.* Am J Hum Genet, 2010. **86**(1): p. 6-22.

31. Ramanan, V.K., et al., *Pathway analysis of genomic data: concepts, methods, and prospects for future development.* Trends Genet, 2012. **28**(7): p. 323-32.

32. Wang, K., M. Li, and H. Hakonarson, *Analysing biological pathways in genome-wide association studies.* Nat Rev Genet, 2010. **11**(12): p. 843-854.

33. Lambert, J.C., et al., *Implication of the immune system in Alzheimer's disease: evidence from genome-wide pathway analysis.* J Alzheimers Dis, 2010. **20**(4): p. 1107-18.

34. Hong, M.-G., et al., *Genome-wide pathway analysis implicates intracellular transmembrane protein transport in Alzheimer disease.* J Hum Genet, 2010. **55**(10): p. 707-709.

35. Ramanan, V.K., et al., *Genome-wide pathway analysis of memory impairment in the Alzheimer's Disease Neuroimaging Initiative (ADNI) cohort implicates gene candidates, canonical pathways, and networks.* Brain Imaging Behav, 2012. **6**(4): p. 634-48.

36. O'Dushlaine, C., et al., *Molecular pathways involved in neuronal cell adhesion and membrane scaffolding contribute to schizophrenia and bipolar disorder susceptibility.* Mol Psychiatry, 2011. **16**(3): p. 286-292.

37. O'Dushlaine, C., et al., *The SNP ratio test: pathway analysis of genome-wide association datasets.* Bioinformatics, 2009. **25**(20): p. 2762-3.

38. Baranzini, S.E., et al., *Pathway and network-based analysis of genome-wide association studies in multiple sclerosis.* Hum Mol Genet, 2009. **18**(11): p. 2078-90.

39. Sawcer, S., et al., *Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis.* Nature, 2011. **476**(7359): p. 214-219.

40. Psychiatric, G.C.B.D.W.G., *Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4.* Nat Genet, 2011. **43**(10): p. 977-83.

41. Hirschhorn, J.N., *Genomewide Association Studies — Illuminating Biologic Pathways.* New England Journal of Medicine, 2009. **360**(17): p. 1699-1701.

42. Penrod, N.M., R. Cowper-Sal-lari, and J.H. Moore, *Systems genetics for drug target discovery.* Trends Pharmacol Sci, 2011. **32**(10): p. 623-30.

43. Bras, J., R. Guerreiro, and J. Hardy, *Use of next-generation sequencing and other whole-genome strategies to dissect neurological disease.* Nat Rev Neurosci, 2012. **13**(7): p. 453-464.

44. Braskie, M.N., J.M. Ringman, and P.M. Thompson, *Neuroimaging measures as endophenotypes in Alzheimer's disease.* Int J Alzheimers Dis, 2011. **2011**: p. 490140.

45. Cruchaga, C., et al., *Cerebrospinal fluid APOE levels: an endophenotype for genetic studies for Alzheimer's disease.* Human Molecular Genetics, 2012.

46. Jack, C.R., Jr., et al., *Shapes of the Trajectories of 5 Major Biomarkers of Alzheimer Disease.* Arch Neurol, 2012. **69**(7): p. 856-867.

47. McKinney, B.A. and N.M. Pajewski, *Six Degrees of Epistasis: Statistical Network Models for GWAS.* Front Genet, 2011. **2**: p. 109.

48. Vidal, M., M.E. Cusick, and A.L. Barabasi, *Interactome networks and human disease.* Cell, 2011. **144**(6): p. 986-98.

49. Schadt, E.E., *Molecular networks as sensors and drivers of common human diseases.* Nature, 2009. **461**(7261): p. 218-23.

50. Jarvik, J. and D. Botstein, *A genetic method for determining the order of events in a biological pathway.* Proc Natl Acad Sci U S A, 1973. **70**(7): p. 2046-50.

51. Holmans, P., *Statistical methods for pathway analysis of genome-wide data for association with complex genetic traits.* Advances in genetics, 2010. **72**: p. 141-179.

52. Marchion, D.C., et al., *BAD phosphorylation determines ovarian cancer chemo-sensitivity and patient survival.* Clinical Cancer Research, 2011.

53. Newman, M.E.J., *Networks: an introduction*. 2010: Oxford University Press.

54. Hu, T., et al., *Characterizing genetic interactions in human disease association studies using statistical epistasis networks.* BMC Bioinformatics, 2011. **12**(1): p. 364.

55. Price, N.D. and I. Shmulevich, *Biochemical and statistical network models for systems biology.* Current Opinion in Biotechnology, 2007. **18**(4): p. 365-370.

56. Shannon, P., et al., *Cytoscape: a software environment for integrated models of biomolecular interaction networks.* Genome Res, 2003. **13**(11): p. 2498-504.

57. Ghosh, S., et al., *Software for systems biology: from tools to integrated platforms.* Nat Rev Genet, 2011. **12**(12): p. 821-832.

58. Thomas, S. and D. Bonchev, *A survey of current software for network analysis in molecular biology.* Hum Genomics, 2010. **4**(5): p. 353-60.

59. Liu, G., et al., *Cell adhesion molecules contribute to Alzheimer's disease: multiple pathway analyses of two genome-wide association studies.* J Neurochem, 2012. **120**(1): p. 190-8.

60. Menashe, I., et al., *Pathway Analysis of Breast Cancer Genome-Wide Association Study Highlights Three Pathways and One Canonical Signaling Cascade.* Cancer Research, 2010. **70**(11): p. 4453-4459.

61. Zhong, H., et al., *Integrating pathway analysis and genetics of gene expression for genome-wide association studies.* Am J Hum Genet, 2010. **86**(4): p. 581-91.

62. Noorbakhsh, F., C.M. Overall, and C. Power, *Deciphering complex mechanisms in neurodegenerative diseases: the advent of systems biology.* Trends Neurosci, 2009. **32**(2): p. 88-100.

63. Hindorff, L.A., et al., *A Catalog of Published Genome-Wide Association Studies.* National Human Genome Research Institute, http://www.genome.gov/gwastudies, 2011.

64. Abatangelo, L., et al., *Comparative study of gene set enrichment methods.* BMC Bioinformatics, 2009. **10**: p. 275.

65. Telander, D.G., *Inflammation and Age-Related Macular Degeneration (AMD).* Seminars in Ophthalmology, 2011. **26**(3): p. 192-197.

66. Eleftherohorinou, H., et al., *Pathway Analysis of GWAS Provides New Insights into Genetic Susceptibility to 3 Inflammatory Diseases.* PLoS One, 2009. **4**(11): p. e8068.

67. Chen, X., et al., *Pathway-based analysis for genome-wide association studies using supervised principal components.* Genetic Epidemiology, 2010. **34**(7): p. 716-724.

68. Zhao, J., et al., *Pathway-based analysis using reduced gene subsets in genome-wide association studies.* BMC Bioinformatics, 2011. **12**: p. 17.

69. Chen, M., J. Cho, and H. Zhao, *Incorporating Biological Pathways via a Markov Random Field Model in Genome-Wide Association Studies.* PLoS Genet, 2011. **7**(4): p. e1001353.

70. Wang, K., M. Li, and M. Bucan, *Pathway-based approaches for analysis of genomewide association studies.* Am J Hum Genet, 2007. **81**(6): p. 1278-83.

71. Askland, K., et al., *Ion channels and schizophrenia: a gene set-based analytic approach to GWAS data for biological hypothesis testing.* Human Genetics, 2011: p. 1-19.

72. Lascorz, J., et al., *Consensus pathways implicated in prognosis of colorectal cancer identified through systematic enrichment analysis of gene expression profiling studies.* PLoS One, 2011. **6**(4): p. e18867.

73. Peng, G., et al., *Gene and pathway-based second-wave analysis of genome-wide association studies.* Eur J Hum Genet, 2010. **18**(1): p. 111-7.

74. Bader, G.D., M.P. Cary, and C. Sander, *Pathguide: a Pathway Resource List.* Nucleic Acids Research, 2006. **34**(suppl 1): p. D504-D506.

75. Camon, E.B., et al., *An evaluation of GO annotation retrieval for BioCreAtIvE and GOA.* BMC Bioinformatics, 2005. **6**.

76. Swaminathan, S., et al., *Amyloid pathway-based candidate gene analysis of [(11)C]PiB-PET in the Alzheimer's Disease Neuroimaging Initiative (ADNI) cohort.* Brain Imaging Behav, 2012. **6**(1): p. 1-15.

77. Sloan, C.D., et al., *Genetic pathway-based hierarchical clustering analysis of older adults with cognitive complaints and amnestic mild cognitive impairment using clinical and neuroimaging phenotypes.* American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 2010. **153B**(5): p. 1060-1069.

78. Zhang, M., et al., *Pathway Analysis for Genome-Wide Association Study of Basal Cell Carcinoma of the Skin.* PLoS One, 2011. **6**(7): p. e22760.

79. Elbers, C.C., et al., *Using genome-wide pathway analysis to unravel the etiology of complex diseases.* Genetic Epidemiology, 2009. **33**(5): p. 419-431.

80. Perry, J.R., et al., *Interrogating type 2 diabetes genome-wide association data using a biological pathway-based approach.* Diabetes, 2009. **58**(6): p. 1463-7.

81. Ballard, D., et al., *Pathway analysis comparison using Crohn's disease genome wide association studies.* BMC Medical Genomics, 2010. **3**(1): p. 25.

82. Yon Rhee, S., et al., *Use and misuse of the gene ontology annotations.* Nat Rev Genet, 2008. **9**(7): p. 509-515.

83. Higareda-Almaraz, J.C., et al., *Proteomic patterns of cervical cancer cell lines, a network perspective.* BMC Syst Biol, 2011. **5**: p. 96.

84. Zhang, S., et al., *GO-Bayes: Gene Ontology-based overrepresentation analysis using a Bayesian approach.* Bioinformatics, 2010. **26**(7): p. 905-11.

85. Alterovitz, G., et al., *Ontology engineering.* Nat Biotech, 2010. **28**(2): p. 128-130.

86. Merico, D., et al., *Enrichment Map: A Network-Based Method for Gene-Set Enrichment Visualization and Interpretation.* PLoS One, 2010. **5**(11): p. e13984.

87. Yaspan, B., et al., *Genetic analysis of biological pathway data through genomic randomization.* Human Genetics, 2011. **129**(5): p. 563-571.

88. Askland, K., C. Read, and J. Moore, *Pathways-based analyses of whole-genome association study data in bipolar disorder reveal genes mediating ion channel activity and synaptic neurotransmission.* Hum Genet, 2009. **125**(1): p. 63-79.

89. Kapranov, P., A.T. Willingham, and T.R. Gingeras, *Genome-wide transcription and the implications for genomic organization.* Nat Rev Genet, 2007. **8**(6): p. 413-23.

90. Portin, P., *The elusive concept of the gene.* Hereditas, 2009. **146**(3): p. 112-7.

91. Edwards, Y.J., et al., *Identifying consensus disease pathways in Parkinson's disease using an integrative systems biology approach.* PLoS One, 2011. **6**(2): p. e16917.

92. Marchini, J. and B. Howie, *Genotype imputation for genome-wide association studies.* Nat Rev Genet, 2010. **11**(7): p. 499-511.

93. Holmans, P., et al., *Gene ontology analysis of GWA study data sets provides insights into the biology of bipolar disorder.* Am J Hum Genet, 2009. **85**(1): p. 13-24.

94. Purcell, S., et al., *PLINK: a tool set for whole-genome association and population-based linkage analyses.* Am J Hum Genet, 2007. **81**(3): p. 559-75.

95. Liu, J.Z., et al., *A Versatile Gene-Based Test for Genome-wide Association Studies.* American Journal of Human Genetics, 2010. **87**(1): p. 139-145.

96. Huang, H., et al., *Gene-Based Tests of Association.* PLoS Genet, 2011. **7**(7): p. e1002177.

97. Gaudet, M.M., et al., *Common genetic variants and modification of penetrance of BRCA2-associated breast cancer.* PLoS Genet, 2010. **6**(10): p. e1001183.

98. Goeman, J.J. and P. Bühlmann, *Analyzing gene expression data in terms of gene sets: methodological issues.* Bioinformatics, 2007. **23**(8): p. 980-987.

99. Fridley, B.L., G.D. Jenkins, and J.M. Biernacka, *Self-contained gene-set analysis of expression data: an evaluation of existing and novel methods.* PLoS One, 2010. **5**(9).

100. Balding, D.J., *A tutorial on statistical methods for population association studies.* Nat Rev Genet, 2006. **7**(10): p. 781-91.

101. Price, A.L., et al., *Principal components analysis corrects for stratification in genome-wide association studies.* Nat Genet, 2006. **38**(8): p. 904-9.

102. Moskvina, V., et al., *Gene-wide analyses of genome-wide association data sets: evidence for multiple common risk alleles for schizophrenia and bipolar disorder and for overlap in genetic risk.* Mol Psychiatry, 2009. **14**(3): p. 252-60.

103. Chen, H. and B. Sharp, *Content-rich biological network constructed by mining PubMed abstracts.* BMC Bioinformatics, 2004. **5**(1): p. 147.

104. Wang, K., et al., *Diverse genome-wide association studies associate the IL12/IL23 pathway with Crohn Disease.* Am J Hum Genet, 2009. **84**(3): p. 399-405.

105. Medina, I., et al., *Gene set-based analysis of polymorphisms: finding pathways or biological processes associated to traits in genome-wide association studies.* Nucleic Acids Res, 2009. **37**(Web Server issue): p. W340-4.

106. Raychaudhuri, S., et al., *Identifying relationships among genomic disease regions: predicting genes at pathogenic SNP associations and rare deletions.* PLoS Genet, 2009. **5**(6): p. e1000534.

107. Chen, L.S., et al., *Insights into colon cancer etiology via a regularized approach to gene set analysis of GWAS data.* Am J Hum Genet, 2010. **86**(6): p. 860-71.

108. Nam, D., et al., *GSA-SNP: a general approach for gene set analysis of polymorphisms.* Nucleic Acids Research, 2010. **38**(suppl 2): p. W749-W754.

109. Subramanian, A., et al., *Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles.* Proc Natl Acad Sci U S A, 2005. **102**(43): p. 15545-50.

110. Holden, M., et al., *GSEA-SNP: applying gene set enrichment analysis to SNP data from genome-wide association studies.* Bioinformatics, 2008. **24**(23): p. 2784-5.

111. Segrè, A.V., et al., *Common Inherited Variation in Mitochondrial Genes Is Not Enriched for Associations with Type 2 Diabetes or Related Glycemic Traits.* PLoS Genet, 2010. **6**(8): p. e1001058.

112. Tintle, N., et al., *Comparing gene set analysis methods on single-nucleotide polymorphism data from Genetic Analysis Workshop 16.* BMC Proc, 2009. **3**(Suppl 7): p. S96.

113. Wang, L., et al., *Gene set analysis of genome-wide association studies: Methodological issues and perspectives.* Genomics, 2011. **98**(1): p. 1-8.

114. Hung, J.-H., et al., *Gene set enrichment analysis: performance evaluation and usage guidelines.* Briefings in Bioinformatics, 2011.

115. Kim, S.-Y. and D. Volsky, *PAGE: Parametric Analysis of Gene Set Enrichment.* BMC Bioinformatics, 2005. **6**(1): p. 144.

116. Berry, M.P.R., et al., *An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis.* Nature, 2010. **466**(7309): p. 973-977.

117. Baranzini, S.E., et al., *Genetic variation influences glutamate concentrations in brains of patients with multiple sclerosis.* Brain, 2010. **133**(9): p. 2603-11.

118. Lee, H.M., et al., *Abnormal networks of immune response-related molecules in bone marrow cells from patients with rheumatoid arthritis as revealed by DNA microarray analysis.* Arthritis Res Ther, 2011. **13**(3): p. R89.

119. Pang, H., M. Hauser, and S. Minvielle, *Pathway-based identification of SNPs predictive of survival.* Eur J Hum Genet, 2011. **19**(6): p. 704-709.

120. Gilman, S.R., et al., *Rare de novo variants associated with autism implicate a large functional network of genes involved in formation and function of synapses.* Neuron, 2011. **70**(5): p. 898-907.

121. Yang, H.-C. and C.-W. Chen, *Region-based and pathway-based QTL mapping using a p-value combination method.* BMC Proc, 2011. **5**(Suppl 9): p. S43.

122. McLean, C.Y., et al., *GREAT improves functional interpretation of cis-regulatory regions.* Nat Biotech, 2010. **28**(5): p. 495-501.

123. Hong, M.-G., et al., *Strategies and issues in the detection of pathway enrichment in genome-wide association studies.* Human Genetics, 2009. **126**(2): p. 289-301.

124. Benjamini, Y. and Y. Hochberg, *Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing.* Journal of the Royal Statistical Society Series B-Methodological, 1995. **57**(1): p. 289-300.

125. Sabatti, C., *Avoiding False Discoveries in Association Studies*, in *Methods in Molecular Biology: Linkage Disequilibrium and Association Mapping*, A.R. Collins, Editor. 2007, Humana Press. p. 195-211.

126. Luo, L., et al., *Genome-wide gene and pathway analysis.* Eur J Hum Genet, 2010. **18**(9): p. 1045-53.

127. Bertram, L., et al., *Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database.* Nat Genet, 2007. **39**(1): p. 17-23.

128. Lill, C.M., et al., *Comprehensive research synopsis and systematic meta-analyses in Parkinson's disease genetics: The PDGene database.* PLoS Genet, 2012. **8**(3): p. e1002548.

129. Zhu, J., et al., *The UCSC Cancer Genomics Browser.* Nat Meth, 2009. **6**(4): p. 239-240.

130. Brown, S.D.M., et al., *The Functional Annotation of Mammalian Genomes: The Challenge of Phenotyping.* Annual Review of Genetics, 2009. **43**(1): p. 305-333.

131. Cowper-Sal·lari, R., et al., *Layers of epistasis: genome-wide regulatory networks and network approaches to genome-wide association studies.* Wiley Interdisciplinary Reviews: Systems Biology and Medicine, 2011. **3**(5): p. 513-526.

132. Bandyopadhyay, S., et al., *Rewiring of Genetic Networks in Response to DNA Damage.* Science, 2010. **330**(6009): p. 1385-1389.

133. Sun, J., et al., *Application of systems biology approach identifies and validates GRB2 as a risk gene for schizophrenia in the Irish Case Control Study of Schizophrenia (ICCSS) sample.* Schizophrenia research, 2011. **125**(2): p. 201-208.

134. Edvardsson, K., et al., *Estrogen receptor beta induces antiinflammatory and antitumorigenic networks in colon cancer cells.* Mol Endocrinol, 2011. **25**(6): p. 969-79.

135. Gui, H., et al., *Comparisons of seven algorithms for pathway analysis using the WTCCC Crohn's Disease dataset.* BMC Research Notes, 2011. **4**(1): p. 386.

136. Gorlov, I.P., et al., *GWAS Meets Microarray: Are the Results of Genome-Wide Association Studies and Gene-Expression Profiling Consistent? Prostate Cancer as an Example.* PLoS One, 2009. **4**(8): p. e6511.

137. Myers, A.J., et al., *A survey of genetic human cortical gene expression.* Nat Genet, 2007. **39**(12): p. 1494-9.

138. Nicolae, D.L., et al., *Trait-Associated SNPs Are More Likely to Be eQTLs: Annotation to Enhance Discovery from GWAS.* PLoS Genet, 2010. **6**(4): p. e1000888.

139. Liang, W.S., et al., *Alzheimer's disease is associated with reduced expression of energy metabolism genes in posterior cingulate neurons.* Proceedings of the National Academy of Sciences, 2008. **105**(11): p. 4441-4446.

140. Le-Niculescu, H., et al., *Convergent functional genomics of genome-wide association data for bipolar disorder: comprehensive identification of candidate genes, pathways and mechanisms.* Am J Med Genet B Neuropsychiatr Genet, 2009. **150B**(2): p. 155-81.

141. Ala-Korpela, M., A.J. Kangas, and M. Inouye, *Genome-wide association studies and systems biology: together at last.* Trends in Genetics, 2011. **27**(12): p. 493-498.

142. Sweatt, J.D., *Mechanisms of Memory, 2nd Edition*. 2009: Academic Press.

143. Tully, T., et al., *Targeting the CREB pathway for memory enhancers.* Nat Rev Drug Discov, 2003. **2**(4): p. 267-77.

144. Reitz, C., C. Brayne, and R. Mayeux, *Epidemiology of Alzheimer disease.* Nat Rev Neurol, 2011. **7**(3): p. 137-52.

145. Papassotiropoulos, A. and D.J. de Quervain, *Genetics of human episodic memory: dealing with complexity.* Trends Cogn Sci, 2011. **15**(9): p. 381-7.

146. Ballard, C., et al., *Alzheimer's disease.* Lancet, 2011. **377**(9770): p. 1019-31.

147. Papassotiropoulos, A., et al., *Common Kibra alleles are associated with human memory performance.* Science, 2006. **314**(5798): p. 475-8.

148. Papassotiropoulos, A., et al., *A genome-wide survey and functional brain imaging study identify CTNNBL1 as a memory-related gene.* Mol Psychiatry, 2011.

149. Harris, S.E. and I.J. Deary, *The genetics of cognitive ability and cognitive ageing in healthy older people.* Trends Cogn Sci, 2011. **15**(9): p. 388-94.

150. Huentelman, M.J., et al., *Calmodulin-binding transcription activator 1 (CAMTA1) alleles predispose human episodic memory performance.* Hum Mol Genet, 2007. **16**(12): p. 1469-77.

151. Milnik, A., et al., *Association of KIBRA with episodic and working memory: A meta-analysis.* Am J Med Genet B Neuropsychiatr Genet, 2012.

152. Papassotiropoulos, A., et al., *A genome-wide survey of human short-term memory.* Mol Psychiatry, 2011. **16**(2): p. 184-192.

153. Lezak, M.D., *Neuropsychological assessment*. 2004: Oxford University Press.

154. Barbeau, E.J., et al., *Extent and Neural Basis of Semantic Memory Impairment in Mild Cognitive Impairment.* Journal of Alzheimer's Disease, 2011.

155. Crane, P.K., et al., *Development and assessment of a composite score for memory in the Alzheimer's Disease Neuroimaging Initiative (ADNI).* Brain Imaging Behav, 2012.

156. Gibbons, L.E., et al., *A composite score for executive functioning, validated in Alzheimer's Disease Neuroimaging Initiative (ADNI) participants with baseline mild cognitive impairment.* Brain Imaging Behav, 2012: p. 1-11.

157. Mukherjee, S., et al., *Genetic architecture of resilience of executive functioning.* Brain Imaging Behav, 2012: p. 1-13.

158. Reeve, B.B., et al., *Psychometric evaluation and calibration of health-related quality of life item banks: plans for the Patient-Reported Outcomes Measurement Information System (PROMIS).* Med Care, 2007. **45**(5 Suppl 1): p. S22-31.

159. Chee Seng, K., et al., *The pursuit of genome-wide association studies: where are we now?* Journal of Human Genetics, 2010. **55**(4): p. 195-206.

160. Han, M.R., et al., *Genome-wide association reveals genetic effects on human Abeta42 and tau protein levels in cerebrospinal fluids: a case control study.* BMC Neurol, 2010. **10**(1): p. 90.

161. Luciano, M., et al., *Whole genome association scan for genetic polymorphisms influencing information processing speed.* Biological Psychology, 2011. **86**(3): p. 193-202.

162. Weiner, M.W., et al., *The Alzheimer's disease neuroimaging initiative: progress report and future plans.* Alzheimers Dement, 2010. **6**(3): p. 202-11 e7.

163. Jack, C.R., et al., *Serial PIB and MRI in normal, mild cognitive impairment and Alzheimer's disease: implications for sequence of pathological events in Alzheimer's disease.* Brain, 2009.

164. Aisen, P.S., et al., *Clinical Core of the Alzheimer's Disease Neuroimaging Initiative: progress and plans.* Alzheimers Dement, 2010. **6**(3): p. 239-46.

165. Folstein, M.F., S.E. Folstein, and P.R. McHugh, *"Mini-mental state". A practical method for grading the cognitive state of patients for the clinician.* J Psychiatr Res, 1975. **12**(3): p. 189-98.

166. Cockrell, J.R. and M.F. Folstein, *Mini-Mental State Examination (MMSE).* Psychopharmacol Bull, 1988. **24**(4): p. 689-92.

167. Mohs, R.C., et al., *Development of cognitive instruments for use in clinical trials of antidementia drugs: additions to the Alzheimer's Disease Assessment Scale that broaden its scope. The Alzheimer's Disease Cooperative Study.* Alzheimer Dis Assoc Disord, 1997. **11 Suppl 2**: p. S13-21.

168. Rey, A., *L'examen clinique en psychologie.* Paris: Presses Universitaires de France, 1964.

169. Wechsler, D., *Wechsler Memory Scale -- Revised.* New York: Psychological Association, 1987.

170. Saykin, A.J., et al., *Alzheimer's Disease Neuroimaging Initiative biomarkers as quantitative phenotypes: Genetics core aims, progress, and plans.* Alzheimers Dement, 2010. **6**(3): p. 265-73.

171. Bertram, L., C.M. Lill, and R.E. Tanzi, *The Genetics of Alzheimer Disease: Back to the Future.* Neuron, 2010. **68**(2): p. 270-281.

172. Strittmatter, W.J., et al., *Apolipoprotein E: high-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease.* Proceedings of the National Academy of Sciences, 1993. **90**(5): p. 1977-1981.

173. Potkin, S.G., et al., *Hippocampal atrophy as a quantitative trait in a genome-wide association study identifying novel susceptibility genes for Alzheimer's disease.* PLoS One, 2009. **4**(8): p. e6501.

174. Shen, L., et al., *Whole genome association study of brain-wide imaging phenotypes for identifying quantitative trait loci in MCI and AD: A study of the ADNI cohort.* Neuroimage, 2010. **53**(3): p. 1051-63.

175. International HapMap, C., *The International HapMap Project.* Nature, 2003. **426**(6968): p. 789-96.

176. Negash, S., et al., *Cognition and neuropathology in aging: multidimensional perspectives from the Rush Religious Orders Study and Rush Memory And Aging Project.* Curr Alzheimer Res, 2011. **8**(4): p. 336-40.

177. Andreano, J.M. and L. Cahill, *Sex influences on the neurobiology of learning and memory.* Learning & Memory, 2009. **16**(4): p. 248-266.

178. Gunstad, J., et al., *Handedness and cognition across the healthy lifespan.* Int J Neurosci, 2007. **117**(4): p. 477-85.

179. Bookheimer, S. and A. Burggren, *APOE-4 Genotype and Neurophysiological Vulnerability to Alzheimer's and Cognitive Aging.* Annual Review of Clinical Psychology, 2009. **5**(1): p. 343-362.

180. Barrett, J.C., et al., *Haploview: analysis and visualization of LD and haplotype maps.* Bioinformatics, 2005. **21**(2): p. 263-5.

181. International Human Genome Sequencing Consortium, *Finishing the euchromatic sequence of the human genome.* Nature, 2004. **431**(7011): p. 931-945.

182. Fridley, B.L. and J.M. Biernacka, *Gene set analysis of SNP data: benefits, challenges, and future directions.* Eur J Hum Genet, 2011. **19**(8): p. 837-43.

183. Kim, S.Y. and D.J. Volsky, *PAGE: parametric analysis of gene set enrichment.* BMC Bioinformatics, 2005. **6**: p. 144.

184. Roses, A.D., et al., *A TOMM40 variable-length polymorphism predicts the age of late-onset Alzheimer's disease.* Pharmacogenomics J, 2010. **10**(5): p. 375-384.

185. Wijsman, E.M., et al., *Genome-Wide Association of Familial Late-Onset Alzheimer's Disease Replicates BIN1 and CLU and Nominates CUGBP2 in Interaction with APOE.* PLoS Genet, 2011. **7**(2): p. e1001308.

186. Johnson, S.C., et al., *The effect of TOMM40 poly-T length on gray matter volume and cognition in middle-aged persons with APOE ε3/ε3 genotype.* Alzheimer's & dementia: the journal of the Alzheimer's Association, 2011. **7**(4): p. 456-465.

187. Sweatt, J.D., *Neuroscience. Creating stable memories.* Science, 2011. **331**(6019): p. 869-70.

188. Collingridge, G.L., et al., *Long-term depression in the CNS.* Nat Rev Neurosci, 2010. **11**(7): p. 459-473.

189. Corvin, A.P., *Neuronal cell adhesion genes: Key players in risk for schizophrenia, bipolar disorder and other neurodevelopmental brain disorders?* Cell Adh Migr, 2010. **4**(4): p. 511-4.

190. Aisa, B., et al., *Altered NCAM Expression Associated with the Cholinergic System in Alzheimer's Disease.* Journal of Alzheimer's Disease, 2010. **20**(2): p. 659-668.

191. Chen, K.P. and F. Dou, *Selective interaction of amyloid precursor protein with different isoforms of neural cell adhesion molecule.* J Mol Neurosci, 2012. **46**(1): p. 203-9.

192. Ho, V.M., J.-A. Lee, and K.C. Martin, *The Cell Biology of Synaptic Plasticity.* Science, 2011. **334**(6056): p. 623-628.

193. Izzi, L. and F. Charron, *Midline axon guidance and human genetic disorders.* Clinical Genetics, 2011. **80**(3): p. 226-234.

194. Arese, M., G. Serini, and F. Bussolino, *Nervous vascular parallels: axon guidance and beyond.* Int J Dev Biol, 2011. **55**(4-5): p. 439-45.

195. Yirmiya, R. and I. Goshen, *Immune modulation of learning, memory, neural plasticity and neurogenesis.* Brain, Behavior, and Immunity, 2011. **25**(2): p. 181-213.

196. Emilio, B., *Microglia: Activation in acute and chronic inflammatory states and in response to cardiovascular dysfunction.* The International Journal of Biochemistry & Cell Biology, 2010. **42**(10): p. 1580-1585.

197. Pawlowski, T.L. and M.J. Huentelman, *Identification of a Common Variant Affecting Human Episodic Memory Performance Using a Pooled Genome-Wide Association Approach: A Case Study of Disease Gene Identification*, J.K. DiStefano, Editor. 2011, Humana Press. p. 261-269.

198. Stein, J.L., et al., *Voxelwise genome-wide association study (vGWAS).* Neuroimage, 2010. **53**(3): p. 1160-74.

199. Stein, J.L., et al., *Genome-wide analysis reveals novel genes influencing temporal lobe structure with relevance to neurodegeneration in Alzheimer's disease.* Neuroimage, 2010. **51**(2): p. 542-54.

200. Hibar, D.P., et al., *Voxelwise gene-wide association study (vGeneWAS): Multivariate gene-based association testing in 731 elderly subjects.* Neuroimage, 2011. **56**(4): p. 1875-1891.

201. Meda, S.A., et al., *A large scale multivariate parallel ICA method reveals novel imaging-genetic relationships for Alzheimer's disease in the ADNI cohort.* Neuroimage, 2012(0).

202. Yu, H.T., et al., *Transcriptional regulation of human FE65, a ligand of Alzheimer's disease amyloid precursor protein, by Sp1.* J Cell Biochem, 2010. **109**(4): p. 782-93.

203. Rossner, S., et al., *Transcriptional and translational regulation of BACE1 expression--implications for Alzheimer's disease.* Prog Neurobiol, 2006. **79**(2): p. 95-111.

204. Santpere, G., et al., *Abnormal Sp1 transcription factor expression in Alzheimer disease and tauopathies.* Neurosci Lett, 2006. **397**(1-2): p. 30-4.

205. Maloney, B., et al., *Functional characterization of three single-nucleotide polymorphisms present in the human APOE promoter sequence: Differential effects in neuronal cells and on DNA-protein interactions.* Am J Med Genet B Neuropsychiatr Genet, 2010. **153B**(1): p. 185-201.

206. Kantarci, K., et al., *APOE modifies the association between Aß load and cognition in cognitively normal older adults.* Neurology, 2012. **78**(4): p. 232-240.

207. Buchman, A.S. and D.A. Bennett, *Amyloid pathology in persons with "normal" cognition.* Neurology, 2012. **78**(4): p. 228-9.

208. Karran, E., M. Mercken, and B. De Strooper, *The amyloid cascade hypothesis for Alzheimer's disease: an appraisal for the development of therapeutics.* Nat Rev Drug Discov, 2011. **10**(9): p. 698-712.

209. Nelson, P.T., et al., *Correlation of Alzheimer disease neuropathologic changes with cognitive status: a review of the literature.* J Neuropathol Exp Neurol, 2012. **71**(5): p. 362-81.

210. Bennett, D.A., et al., *Neuropathology of older persons without cognitive impairment from two community-based studies.* Neurology, 2006. **66**(12): p. 1837-44.

211. Kim, J., J.M. Basak, and D.M. Holtzman, *The role of apolipoprotein E in Alzheimer's disease.* Neuron, 2009. **63**(3): p. 287-303.

212. Furst, A.J. and G.A. Kerchner, *From Alois to Amyvid: Seeing Alzheimer disease.* Neurology, 2012.

213. Clark, C.M., et al., *Use of florbetapir-PET for imaging beta-amyloid pathology.* JAMA, 2011. **305**(3): p. 275-83.

214. Clark, C.M., et al., *Cerebral PET with florbetapir compared with neuropathology at autopsy for detection of neuritic amyloid-β plaques: a prospective cohort study.* The Lancet Neurology, 2012. **11**(8): p. 669-678.

215. Jagust, W.J., et al., *The Alzheimer's Disease Neuroimaging Initiative positron emission tomography core.* Alzheimers Dement, 2010. **6**(3): p. 221-9.

216. Howie, B., et al., *Fast and accurate genotype imputation in genome-wide association studies through pre-phasing.* Nat Genet, 2012. **44**(8): p. 955-9.

217. Stein, J.L., et al., *Identification of common variants associated with human hippocampal and intracranial volumes.* Nat Genet, 2012. **44**(5): p. 552-61.

218. Pe'er, I., et al., *Estimation of the multiple testing burden for genomewide association studies of nearly all common variants.* Genet Epidemiol, 2008. **32**(4): p. 381-5.

219. Johnson, A.D., et al., *SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap.* Bioinformatics, 2008. **24**(24): p. 2938-9.

220. Pruim, R.J., et al., *LocusZoom: regional visualization of genome-wide association scan results.* Bioinformatics, 2010. **26**(18): p. 2336-7.

221. Li, M.-X., et al., *GATES: A Rapid and Powerful Gene-Based Association Test Using Extended Simes Procedure.* American Journal of Human Genetics, 2011. **88**(3): p. 283-293.

222. Castellano, J.M., et al., *Human apoE isoforms differentially regulate brain amyloid-beta peptide clearance.* Sci Transl Med, 2011. **3**(89): p. 89ra57.

223. Holtzman, D.M., J.C. Morris, and A.M. Goate, *Alzheimer's disease: the challenge of the second century.* Sci Transl Med, 2011. **3**(77): p. 77sr1.

224. Koffie, R.M., et al., *Apolipoprotein E4 effects in Alzheimer's disease are mediated by synaptotoxic oligomeric amyloid-beta.* Brain, 2012. **135**(Pt 7): p. 2155-68.

225. Lehmann, D.J., C. Johnston, and A.D. Smith, *Synergy Between the Genes for Butyrylcholinesterase K Variant and Apolipoprotein E4 in Late-Onset Confirmed Alzheimer's Disease.* Human Molecular Genetics, 1997. **6**(11): p. 1933-1936.

226. Darvesh, S., D.A. Hopkins, and C. Geula, *Neurobiology of butyrylcholinesterase.* Nat Rev Neurosci, 2003. **4**(2): p. 131-8.

227. Guillozet, A.L., et al., *Butyrylcholinesterase in the life cycle of amyloid plaques.* Ann Neurol, 1997. **42**(6): p. 909-18.

228. Lane, R., et al., *Synergistic effect of apolipoprotein E epsilon4 and butyrylcholinesterase K-variant on progression from mild cognitive impairment to Alzheimer's disease.* Pharmacogenet Genomics, 2008. **18**(4): p. 289-98.

229. Tasker, A., et al., *Butyrylcholinesterase K variant associated with higher enzyme activity in the temporal cortex of elderly patients.* Neurosci Lett, 2008. **442**(3): p. 297-9.

230. Lane, R.M. and Y. He, *Butyrylcholinesterase genotype and gender influence Alzheimer's disease phenotype.* Alzheimers Dement, 2012.

231. Sabbagh, M.N., et al., *Do cholinergic therapies have disease-modifying effects in Alzheimer's disease?* Alzheimer's & dementia : the journal of the Alzheimer's Association, 2006. **2**(2): p. 118-125.

232. Darvesh, S., et al., *Butyrylcholinesterase is associated with beta-amyloid plaques in the transgenic APPSWE/PSEN1dE9 mouse model of Alzheimer disease.* J Neuropathol Exp Neurol, 2012. **71**(1): p. 2-14.

233. Darreh-Shori, T., et al., *Differential levels of apolipoprotein E and butyrylcholinesterase show strong association with pathological signs of Alzheimer's disease in the brain in vivo.* Neurobiology of Aging, 2011. **32**(12): p. 2320.e15-2320.e32.

234. Darreh-Shori, T., et al., *The apolipoprotein E ε4 allele plays pathological roles in AD through high protein expression and interaction with butyrylcholinesterase.* Neurobiology of Aging, 2011. **32**(7): p. 1236-1248.

235. Lane, R.M. and Y. He, *Emerging hypotheses regarding the influences of butyrylcholinesterase-K variant, APOE ε4, and hyperhomocysteinemia in neurodegenerative dementias.* Medical hypotheses, 2009. **73**(2): p. 230-250.

236. Iwasaki, T., et al., *Serum Butyrylcholinesterase is Strongly Associated with Adiposity, the Serum Lipid Profile and Insulin Resistance.* Internal Medicine, 2007. **46**(19): p. 1633-1639.

237. Meigs, J.B., et al., *Genome-wide association with diabetes-related traits in the Framingham Heart Study.* BMC Med Genet, 2007. **8 Suppl 1**: p. S16.

238. Benyamin, B., et al., *GWAS of butyrylcholinesterase activity identifies four novel loci, independent effects within BCHE and secondary associations with metabolic risk factors.* Hum Mol Genet, 2011. **20**(22): p. 4504-14.

239. Wang, L., et al., *Epidermal growth factor receptor is a preferred target for treating amyloid-beta-induced memory loss.* Proc Natl Acad Sci U S A, 2012. **109**(41): p. 16743-8.

240. Costa, R.M., T. Honjo, and A.J. Silva, *Learning and memory deficits in Notch mutant mice.* Curr Biol, 2003. **13**(15): p. 1348-54.

241. Donmez, G., et al., *SIRT1 suppresses beta-amyloid production by activating the alpha-secretase gene ADAM10.* Cell, 2010. **142**(2): p. 320-32.

242. Cai, Z. and A. Ratka, *Opioid system and Alzheimer's disease.* Neuromolecular Med, 2012. **14**(2): p. 91-111.

243. Durakoglugil, M.S., et al., *Reelin signaling antagonizes beta-amyloid at the synapse.* Proc Natl Acad Sci U S A, 2009. **106**(37): p. 15938-43.

244. Bamberger, M.E., et al., *A cell surface receptor complex for fibrillar beta-amyloid mediates microglial activation.* J Neurosci, 2003. **23**(7): p. 2665-74.

245. Bomfim, T.R., et al., *An anti-diabetes agent protects the mouse brain from defective insulin signaling caused by Alzheimer's disease- associated Abeta oligomers.* J Clin Invest, 2012. **122**(4): p. 1339-53.

246. Ikonomovic, M.D., et al., *Post-mortem correlates of in vivo PiB-PET amyloid imaging in a typical case of Alzheimer's disease.* Brain, 2008. **131**(Pt 6): p. 1630-45.

247. Ramanan, V.K., et al., *APOE and BCHE as modulators of cerebral amyloid deposition: a florbetapir PET genome-wide association study.* Mol Psychiatry, 2013.

248. Bertram, L. and R.E. Tanzi, *The genetic epidemiology of neurodegenerative disease.* J Clin Invest, 2005. **115**(6): p. 1449-1457.

249. Gandhi, S. and N.W. Wood, *Genome-wide association studies: the key to unlocking neurodegeneration?* Nat Neurosci, 2010. **13**(7): p. 789-794.

250. Citron, B.A., et al., *Transcription factor Sp1 dysregulation in Alzheimer's disease.* J Neurosci Res, 2008. **86**(11): p. 2499-504.

251. Chu, J., J.M. Zhuo, and D. Pratico, *Transcriptional regulation of beta-secretase-1 by 12/15-lipoxygenase results in enhanced amyloidogenesis and cognitive impairments.* Ann Neurol, 2012. **71**(1): p. 57-67.

252. Karin, M., Z. Liu, and E. Zandi, *AP-1 function and regulation.* Curr Opin Cell Biol, 1997. **9**(2): p. 240-6.

253. Calon, F., et al., *Molecular basis of levodopa-induced dyskinesias.* Ann Neurol, 2000. **47**(4 Suppl 1): p. S70-8.

254. Andersson, M., C. Konradi, and M.A. Cenci, *cAMP response element-binding protein is required for dopamine-dependent gene expression in the intact but not the dopamine-denervated striatum.* J Neurosci, 2001. **21**(24): p. 9930-43.

255. Jegga, A.G., et al., *Systems biology of the autophagy-lysosomal pathway.* Autophagy, 2011. **7**(5): p. 477-89.

256. Fahrig, T., I. Gerlach, and E. Horvath, *A synthetic derivative of the natural product rocaglaol is a potent inhibitor of cytokine-mediated signaling and shows neuroprotective activity in vitro and in animal models of Parkinson's disease and traumatic brain injury.* Mol Pharmacol, 2005. **67**(5): p. 1544-55.

257. Perez-Cadahia, B., B. Drobic, and J.R. Davie, *Activation and function of immediate-early genes in the nervous system.* Biochem Cell Biol, 2011. **89**(1): p. 61-73.

258. Gomez Ravetti, M., et al., *Uncovering molecular biomarkers that correlate cognitive decline with the changes of hippocampus' gene expression profiles in Alzheimer's disease.* PLoS One, 2010. **5**(4): p. e10153.

259. Lu, Y., et al., *Early growth response 1 (Egr-1) regulates phosphorylation of microtubule-associated protein tau in mammalian brain.* J Biol Chem, 2011. **286**(23): p. 20569-81.

260. Lindersson, E.K., et al., *alpha-Synuclein filaments bind the transcriptional regulator HMGB-1.* Neuroreport, 2004. **15**(18): p. 2735-9.

261. Takata, K., et al., *Role of high mobility group protein-1 (HMG1) in amyloid-beta homeostasis.* Biochem Biophys Res Commun, 2003. **301**(3): p. 699-703.

262. Takata, K., et al., *Microglial Amyloid-beta1-40 Phagocytosis Dysfunction Is Caused by High-Mobility Group Box Protein-1: Implications for the Pathological Progression of Alzheimer's Disease.* Int J Alzheimers Dis, 2012. **2012**: p. 685739.

263. Scaffidi, P., T. Misteli, and M.E. Bianchi, *Release of chromatin protein HMGB1 by necrotic cells triggers inflammation.* Nature, 2002. **418**(6894): p. 191-5.

264. Gao, H.M., et al., *HMGB1 acts on microglia Mac1 to mediate chronic neuroinflammation that drives progressive neurodegeneration.* J Neurosci, 2011. **31**(3): p. 1081-92.

265. Clifford, D.B., et al., *CSF biomarkers of Alzheimer disease in HIV-associated neurologic disease.* Neurology, 2009. **73**(23): p. 1982-7.

266. Giunta, B., et al., *HIV-1 Tat contributes to Alzheimer's disease-like pathology in PSAPP mice.* Int J Clin Exp Pathol, 2009. **2**(5): p. 433-43.

267. Hickman, S.E., E.K. Allison, and J. El Khoury, *Microglial dysfunction and defective beta-amyloid clearance pathways in aging Alzheimer's disease mice.* J Neurosci, 2008. **28**(33): p. 8354-60.

268. Tarassishin, L., et al., *Interferon regulatory factor 3 inhibits astrocyte inflammatory gene expression through suppression of the proinflammatory miR-155 and miR-155*.* Glia, 2011. **59**(12): p. 1911-22.

269. Chakrabarty, P., et al., *Interferon-gamma induces progressive nigrostriatal degeneration and basal ganglia calcification.* Nat Neurosci, 2011. **14**(6): p. 694-6.

270. Soares, H.D., et al., *Plasma biomarkers associated with the apolipoprotein E genotype and Alzheimer disease.* Arch Neurol, 2012. **69**(10): p. 1310-7.

271. Barcia, C., et al., *IFN-gamma signaling, with the synergistic contribution of TNF-alpha, mediates cell specific microglial and astroglial activation in experimental models of Parkinson's disease.* Cell Death Dis, 2011. **2**: p. e142.

272. Jaeger, S. and C.U. Pietrzik, *Functional role of lipoprotein receptors in Alzheimer's disease.* Curr Alzheimer Res, 2008. **5**(1): p. 15-25.

273. Waldron, E., et al., *LRP1 modulates APP trafficking along early compartments of the secretory pathway.* Neurobiol Dis, 2008. **31**(2): p. 188-97.

274. Liu, Q., et al., *Neuronal LRP1 knockout in adult mice leads to impaired brain lipid metabolism and progressive, age-dependent synapse loss and neurodegeneration.* J Neurosci, 2010. **30**(50): p. 17068-78.

275. Sloan, C.D., et al., *Genetic pathway-based hierarchical clustering analysis of older adults with cognitive complaints and amnestic mild cognitive impairment using clinical and neuroimaging phenotypes.* Am J Med Genet B Neuropsychiatr Genet, 2010. **153B**(5): p. 1060-9.

276. Peregrin-Alvarez, J.M., C. Sanford, and J. Parkinson, *The conservation and evolutionary modularity of metabolism.* Genome Biol, 2009. **10**(6): p. R63.

277. Ayalew, M., et al., *Convergent functional genomics of schizophrenia: from comprehensive understanding to genetic risk prediction.* Mol Psychiatry, 2012. **17**(9): p. 887-905.

CURRICULUM VITAE

Vijay K Ramanan

Education

2006-2013 **Indiana University School of Medicine, Indianapolis, IN**
Medical Scientist (MD/PhD) Training Program
PhD, Medical and Molecular Genetics (June 2013)
GPA: 3.93
Thesis: "Pathways to dementia: genetic predictors of cognitive and brain imaging endophenotypes in Alzheimer's disease"
Advisor: Dr. Andrew J. Saykin, PsyD
MD (expected May 2015)

2002-2006 **University of Notre Dame, South Bend, IN**
Glynn Scholars Honors Program
BS, Honors Mathematics; Minor, Philosophy
GPA: 3.79
Honors Thesis: "Analysis of gene expression during light-induced retinal degeneration and regeneration in albino zebrafish"
Advisor: Dr. David R. Hyde, PhD

Research Experience

2011-2013 **Indiana University School of Medicine, Indianapolis, IN**
PhD student (Advisor: Dr. Andrew J. Saykin, PsyD)
1. Concepts and methods for biological pathway analysis of genomic data
2. Biological pathway analysis of memory performance in Alzheimer's disease
3. Environmental and genetic predictors of memory in the general population
4. Genome-wide association study of brain amyloid load in Alzheimer's disease
5. Structural MRI study of hippocampal subfield volumes in Alzheimer's disease
6. Biological pathways implicated across neurodegenerative diseases

2005 **National Cancer Institute, National Institutes of Health, Bethesda, MD**
Student Research Fellow (Advisor: Dr. John C. Morris, MD)
1. Validation of rat models for studying adenoviral cancer gene therapies

2003-2006 **University of Notre Dame, South Bend, IN**
Undergraduate research (Advisor: Dr. David R. Hyde, PhD)
1. Gene expression studies of zebrafish retinal degeneration and regeneration

Clinical Experience

2011-2013 **Indiana University School of Medicine, Indianapolis, IN**
Indiana Memory and Aging Study Clinical Consensus conferences
Advisor: Dr. Andrew J. Saykin, PsyD

2010 **Riley Hospital for Children, Indianapolis, IN**
Medical student rotation
Advisors: Drs. David D. Weaver, MD and Wilfredo Torres-Martinez, MD

2005 **National Cancer Institute Hematology/Oncology clinic**
Student participant in physician rounds
Advisor: Dr. John C. Morris, MD

Other Professional Experience, Honors, Memberships, and Service

2013 Travel fellowship, Alzheimer's Association International Conference, Boston, MA
2013 Panel participant, RCR educational session on authorship, Indiana University
2012 NINDS/AUPN/ANA Combining Clinical/Research Careers Symposium, Washington, DC
2012 Travel fellowship, Alzheimer's Association International Conference, Vancouver, BC
2012 Travel award, Campbell-Klatte Research Day, Department of Radiology and Imaging Sciences, Indiana University School of Medicine
2011 Student member, American Society for Human Genetics
2010-12 Student member, AAAS/Science Excellence Program
2010-11 Student member, Curriculum Reform Committee, Indiana University School of Medicine
2008- MD/PhD student committee, Indiana University School of Medicine
2006 Best Senior Honors Thesis, College of Science, University of Notre Dame
2006 Irish Clover Award for outstanding service, University of Notre Dame
2005-06 Student representative, University Academic Council, University of Notre Dame
2005-06 Student member, Alpha Epsilon Delta pre-medical honor society
2005 Nominee for Rhodes and Goldwater Scholarships, University of Notre Dame
2005 Student teaching assistant, Department of Biological Sciences, University of Notre Dame
2004-05 Co-chair, University Committee on Course Evaluation Reform, University of Notre Dame
2004-05 Research fellowship, Glynn Scholars Honors Program, University of Notre Dame
2002-06 Tutor for undergraduate students, University of Notre Dame
2002-06 Lilly Endowment Scholarship awarding full tuition for undergraduate studies
2002 Notre Dame Scholar, awarded to top 5% of incoming class
2002 *Indianapolis Star* Indiana Academic All-Star
2002 Valedictorian, Clay High School, South Bend, IN
2001-02 Student Body and National Honor Society President, Clay High School, South Bend, IN
1998-2002 Class President, Clay High School, South Bend, IN
1995 Participant in Scripps-Howard National Spelling Bee

Peer-Reviewed Publications

Manuscripts

1. **Ramanan VK**, Risacher SL, Nho K, Kim S, Swaminathan S, Shen L, Foroud TM, Hakonarson H, Huentelman MJ, Aisen PS, Petersen RC, Green RC, Jack CR, Koeppe RA, Jagust WJ, Weiner MW, and Saykin AJ, for the Alzheimer's Disease Neuroimaging Initiative (ADNI). *APOE* and *BCHE* as modulators of cerebral amyloid deposition: a florbetapir PET genome-wide association study. Molecular Psychiatry. 2013 (Advance Online Publication).
2. Nho K, Corneveaux JJ, Kim S, Lin H, Risacher SL, Shen L, Swaminathan S, **Ramanan VK**, Liu Y, Foroud TM, Inlow MH, Siniard AL, Reiman RA, Aisen PS, Petersen RC, Green RC, Jack CR, Weiner MW, Baldwin CT, Lunetta K, Farrer LA, for the Multi-Institutional Research on Alzheimer's Genetic Epidemiology (MIRAGE) Study, Furney SJ, Lovestone S, Simmons A, Mecocci P, Vellas B, Tsolaki M, Kloszewska I, Soininen H, for the AddNeuroMed Consortium, McDonald BC, Farlow MR, Ghetti B, for the Indiana Memory and Aging Study, Huentelman MJ, Saykin AJ, for the Alzheimer's Disease Neuroimaging Initiative (ADNI). Whole-exome sequencing and imaging genetics identify functional variants for rate of change in hippocampal volume in Mild Cognitive Impairment. Molecular Psychiatry. 2013 (In Press).
3. **Ramanan VK**, Kim S, Holohan K, Shen L, Nho K, Risacher SL, Foroud TM, Mukherjee S, Crane PK, Aisen PS, Petersen RC, Weiner MW, and Saykin AJ, for the Alzheimer's Disease Neuroimaging Initiative (ADNI). Genome-wide pathway analysis of memory impairment in the Alzheimer's Disease Neuroimaging Initiative (ADNI) cohort implicates gene candidates, canonical pathways, and networks. Brain Imaging and Behavior. 2012; 6(4):634-48.
4. **Ramanan VK**, Shen L, Moore JH, and Saykin AJ. Pathway analysis of genomic data: concepts, methods, and prospects for future development. Trends in Genetics. 2012; 28(7):323-32.
5. Kassen SC, **Ramanan VK**, Montgomery JE, Burket CT, Liu CG, Vihtelic TS, and Hyde DR. Time course analysis of gene expression during light-induced photoreceptor cell death and regeneration in albino zebrafish. Developmental Neurobiology. 2007; 67(8): 1009-31.
6. Steel JC, Morrison BJ, Mannan P, Abu-Asab MS, Wildner O, Miles BK, Yim KC, **Ramanan VK**, Prince GA, and Morris JC. Immunocompetent syngeneic cotton rat tumor models for the assessment of replication-competent oncolytic adenovirus. Virology. 2007; 369(1):131-42.

Abstracts

1. **Ramanan VK**, Nho K, Shen L, Risacher SL, Kim S, Foroud TM, Bennett DA, Saykin AJ, for the Alzheimer's Disease Neuroimaging Initiative (ADNI). Predictors of memory performance in a large stratified random population sample of older Americans: clinical, demographic, lifestyle, and genetic factors. Alzheimer's Association International Conference 2013. Boston, MA.
2. Shen L, Thompson PM, Potkin SG, Stone DJ, Kim S, Nho K, **Ramanan VK**, Green RC, Foroud TM, Farrer LA, Moore JH, Bertram L, Weiner MW, Saykin AJ, for the Alzheimer's Disease Neuroimaging Initiative (ADNI). A review of published genetic studies using ADNI multimodality quantitative phenotypes: MRI, PET, fluid biomarkers, cognition, and clinical status. Alzheimer's Association International Conference 2013. Boston, MA.
3. Nho K, Corneveaux JJ, Kim S, Lin H, Risacher SL, Shen L, Swaminathan S, **Ramanan VK**, Liu Y, Foroud TM, Inlow MH, Siniard AL, Reiman RA, Aisen PS, Petersen RC, Green RC,

Jack CR, Weiner MW, Baldwin CT, Lunetta K, Farrer LA, for the Multi-Institutional Research on Alzheimer's Genetic Epidemiology (MIRAGE) Study, Furney SJ, Lovestone S, Simmons A, Mecocci P, Vellas B, Tsolaki M, Kloszewska I, Soininen H, for the AddNeuroMed Consortium, McDonald BC, Farlow MR, Ghetti B, for the Indiana Memory and Aging Study, Huentelman MJ, Saykin AJ, for the Alzheimer's Disease Neuroimaging Initiative (ADNI). Protective variant for rate of hippocampal volume loss identified by whole exome sequencing in APOE ε3/ε3 males with MCI. Alzheimer's Association International Conference 2013. Boston, MA.

4. Kim S, Nho K, Shen L, Kling MA, Han X, Zhu H, Sullivan PM, Arnold SE, Risacher SL, **Ramanan VK**, Doraiswamy PM, Trojanowski JQ, Kaddurah-Daouk R, Saykin AJ, for the Alzheimer's Disease Neuroimaging Initiative (ADNI). Targeted lipidomic pathway-guided genetic association with Alzheimer's disease-relevant endophenotypes. Alzheimer's Association International Conference 2013. Boston, MA.
5. **Ramanan VK**, Shen L, West JD, Kim S, Risacher SL, Nho K, Wang Y, Wan J, Foroud TM, Farrer LA, Saykin AJ. Genome-wide pathway analysis in multiple cohorts implicates histone genes and inflammation in hippocampal atrophy. American Society for Human Genetics Annual Meeting 2012. San Francisco, CA.
6. Holohan KN, Risacher SL, Swaminathan S, West JD, Inlow M, Conroy SK, **Ramanan VK**, Foroud TM, Shen L, Saykin AJ. Neural cell adhesion gene variation and brain morphometry in Alzheimer's disease. American Society for Human Genetics Annual Meeting 2012. San Francisco, CA.
7. **Ramanan VK**, Kim S, Holohan KN, Shen L, Nho K, Risacher SL, Foroud TM, Mukherjee S, Crane PK, Aisen PS, Petersen RC, Weiner MW, Saykin AJ, for the Alzheimer's Disease Neuroimaging Initiative (ADNI). Common gene networks underlying memory impairment: a comparison of pathway analysis methods in the Alzheimer's Disease Neuroimaging Initiative (ADNI) cohort. Alzheimer's Association International Conference 2012. Vancouver, BC.
8. **Ramanan VK**, Shen L, Risacher SL, Kim S, Boddu M, West JD, Wan J, Wang Y, Crane PK, Jack CR,Weiner MW, Saykin AJ, for the Alzheimer's Disease Neuroimaging Initiative (ADNI). Hippocampal subfield atrophy in the Alzheimer's Disease Neuroimaging Initiative (ADNI) cohort: relationships to diagnosis and memory impairment. Alzheimer's Association International Conference 2012. Vancouver, BC.
9. Wang Y, West JD, MaGee TR, McDonald BC, Risacher SL, Shen L, **Ramanan VK**, Kim S, O'Neill DP, Farlow MR, Ghetti B, Saykin AJ. Hippocampal subfield atrophy on 3T MRI in prodromal Alzheimer's disease and older adults with cognitive complaints. Alzheimer's Association International Conference 2012. Vancouver, BC.
10. **Ramanan VK**, Kim S, Holohan KN, Shen L, Nho K, Risacher SL, Foroud TM, Mukherjee S, Crane PK, Aisen PS, Petersen RC, Weiner MW, Saykin AJ, for the Alzheimer's Disease Neuroimaging Initiative (ADNI). Genome-wide pathway analysis of memory impairment in the Alzheimer's Disease Neuroimaging Initiative (ADNI) cohort. National MD/PhD Student Conference 2012. Keystone, CO.
11. Wan J, **Ramanan VK**, Kim S, West JD, Boddu M, Risacher SL, Fang S, Saykin AJ, and Shen L. Classification of AD and MCI using hippocampal subfields: an ADNI study. 18th Annual Meeting of the Organization for Human Brain Mapping 2012. Beijing, China.

Research Support

1. Indiana University Medical Scientist Training Program grant NIGMS GM077229-02.