Using Genetic Diversity to Understand Susceptibility to Cognitive Decline in Aging and Alzheimer’s Disease

Sarah M. Neuner
University of Tennessee Health Science Center

Follow this and additional works at: https://dc.uthsc.edu/dissertations
Part of the Genetics Commons, and the Neurosciences Commons

Recommended Citation
Using Genetic Diversity to Understand Susceptibility to Cognitive Decline in Aging and Alzheimer's Disease

Abstract
An individual's genetic makeup plays an important role in determining susceptibility to cognitive aging and transition to dementia such as Alzheimer's disease (AD). Identifying the specific genetic variants that contribute to cognitive aging and AD may aid in early diagnosis of at-risk patients, as well as identify novel therapeutics targets to treat or prevent development of symptoms. Challenges to identifying these specific genes in human studies include complex genetics, difficulty in controlling environmental factors, and limited access to human brain tissue. Here, we turned to genetically diverse mice from the BXD genetic reference panel (GRP) to overcome some of the barriers traditionally associated with human studies. Using a forward genetics screen, we first identified and validated the gene heterochromatin protein I binding protein 3 (H3BP3) as a novel modulator of normal cognitive aging. We then demonstrated that targeted knockdown of H3BP3 in the hippocampus by 50-75% was sufficient to induce cognitive deficits and transcriptional changes reminiscent of those observed in aging and AD, namely an increase in inflammatory pathways and decrease in neuronal and synaptically-localized transcripts. We also show H3BP3 is a translationally relevant target, as transcriptional changes induced by our targeted knockdown significantly overlapped those observed in the aging human brain. In addition, H3BP3 itself was decreased in the hippocampus of cognitively impaired aging humans. In summary, our results suggest therapeutics designed to target either H3BP3 or its downstream effectors may be useful promoting cognitive longevity. We next expanded on our findings that the BXDs were variably susceptible to cognitive aging and combined the BXD GRP with a well-established transgenic mouse model of AD harboring 5 familial AD mutations, the 5XFAD model. The resulting panel, which we termed the AD-BXDs, consists of genetically diverse yet isogenic F1 mice that all harbor the same high-risk human AD mutations but who differ across the remainder of their genome. We first showed that genetic variation profoundly modified the impact of human AD mutations on both cognitive and pathological phenotypes. We then validated this complex AD model by demonstrating high degrees of genetic, transcriptomic, and phenotypic overlap with human AD. Genetic mapping was used to identify novel genomic loci that modified susceptibility or resilience to cognitive and pathological symptoms of AD. Finally, we used transcriptome profiling to identify gene networks present in the pre-symptomatic mouse brain that predict cognitive performance at an advanced age. Together, the candidates identified through these analyses highlight new potential drivers of susceptibility or resilience to AD and contribute significantly to our understanding of early, potentially causal disease mechanisms. In summary, work here highlights the utility of genetically diverse mice to elucidate mechanisms underlying complex human disease, namely cognitive aging and AD. In addition, we developed a novel AD mouse population as an innovative and reproducible resource for the study of mechanisms underlying AD. Data presented here provides convincing evidence that preclinical models incorporating genetic diversity may better translate to human disease. Due to the reproducible nature of the BXDs and resulting AD-BXDs, this approach creates substantial opportunities to develop improved models of human aging and AD as well as develop a better understanding of precise mechanisms underlying disease. Together, these resources may ultimately enable precision medicine approaches across a diverse population. Finally, our experimental design is likely to be broadly applicable to mouse models of human disease that incorporate a dominantly inherited high-risk genotype in the form of a transgene or other genetic perturbation, enhancing the general utility of results reported here.

Document Type
Dissertation
Degree Name
Doctor of Philosophy (PhD)

Program
Biomedical Sciences

Research Advisor
Robert W. Williams, PhD.

Keywords
aging, Alzheimer's disease, cognitive decline, genetics, mouse models

Subject Categories
Genetics | Medicine and Health Sciences | Neurosciences

This dissertation is available at UTHSC Digital Commons: https://dc.uthsc.edu/dissertations/485
Using Genetic Diversity to Understand Susceptibility to Cognitive Decline in Aging and Alzheimer’s Disease

A Dissertation
Presented for
The Graduate Studies Council
The University of Tennessee
Health Science Center

In Partial Fulfillment
Of the Requirements for the Degree
Doctor of Philosophy
From The University of Tennessee

By
Sarah M. Neuner
May 2019
Chapter 3 © 2019 Elsevier, Inc.
All other material © 2019 by Sarah M. Neuner.
All rights reserved.
DEDICATION

I dedicate this body of work to my family, without whom I would not be where I am today. Thank you for your unwavering support and for being a constant source of positivity and encouragement.
ACKNOWLEDGEMENTS

Completion of my PhD would not have been possible without the amazing support of my mentor, Dr. Catherine Kaczorowski. Catherine has truly been both a scientific and personal role model to me every day of this journey. A million thank yous would not be enough to acknowledge how much your support has meant to me through these last six years. Thank you for teaching me, thank you for encouraging me, thank you for challenging me, and thank you for believing in me.

To my dissertation committee: Dr. Kristen O’Connell, Dr. Robert Williams, Dr. William Armstrong, Dr. Dave Bridges, and Dr. Khyobeni Mozhui, I thank you for your continuous support and advice. I have learned so much from each and every one of you and together, you have made me a better person and scientist. To Dr. Rob Williams in particular, thank you for taking me and Catherine under your “systems genetics wing” and teaching us all you know about mouse genetics. Thank you for officially becoming my advisor during our transition to The Jackson Laboratory, which allowed me to remain at UTHSC as a graduate student and complete my PhD. To Dr. Kristen O’Connell, my unofficial third mentor, thank you for everything you have done for me and these projects. Thank you for the hours spent patiently troubleshooting experiments, the last-minute finding and/or lending of supplies, and the hours spent in and out of lab teaching me about science (and life).

I am grateful to the University of Tennessee for admitting me into their PhD program. I knew from my very first interview that UT was a special place, and I was not disappointed. I would particularly like to thank the Neuroscience program, the directors and faculty, the administrative staff, and the other Neuroscience track students for their help and support. Thank you also to the UTHSC Center for Integrative and Translational Genomics for making the initial studies outlined here possible. Thank you to the Neuroscience Institute and the College of Graduate Health Sciences for providing me with multiple travel awards so that I could attend both national and international conferences during my time as a graduate student.

I would also like to thank The Jackson Laboratory, who welcomed me with open arms in 2016 and made the at-times overwhelming move from Memphis to Bar Harbor more manageable. Thank you for teaching me about mouse genetics, for introducing me to the wonderful world of outreach and education, for giving me lifelong friends, and a place I will always call home. Thank you to Alex Fine, my first friend at JAX, and to Carrie Cowan, without whom my graduate experience at JAX would likely have been completely different. Thank you both for always making me feel included. A special thank you to Nick Tolman and the beautiful Fiona, who always keep me laughing and keep me grounded.

I would not have been able to complete the projects outline here without the incredible group of people I have been lucky enough to be surrounded by throughout the last six years. To the entire Kaczorowski and O’Connell labs, past and present, thank you
for your constant willingness to provide support and feedback on these projects and papers, for being an extra set(s) of hands when I needed them most, and for the numerous laughs throughout the years, both in and out of lab. In addition, I would like to thank the members of the Williams laboratory for welcoming me into the world of systems genetics, particularly Dr. Megan Mulligan and Jesse Ingels for their brilliance and patience. I would also like to acknowledge everyone who collaborated with our group on making this project a reality, particularly Dr. Matthew Huentelman and his outstanding group at The Translational Genomics Research Institute as well as Dr. Vivek Philip, Dr. Ji-Gang Zhang, Dr. Kwangbom Choi, and The Jackson Laboratory Genome Technologies Core for all of their assistance with RNA sequencing, including library prep and data processing, alignment, and analysis.

Thank you to the National Institutes on Aging, who believed in this project from the beginning and funded my F31 Ruth L. Kirschstein National Research Service Award Individual Predoctoral Fellowship in 2015. They have also provided our lab with a number of additional grants in the form of R01s to Dr. Kaczorowski (R01AG057914 and R01AG054180), and for that we will be forever thankful. Thank you to the other funding agencies who made this work possible, including the American Federation for Aging Research, the BrightFocus Foundation (A2016397S), the UTHSC Neuroscience Institute, the UTHSC Center for Integrative and Translational Genomics, and The Jackson Laboratory.

Finally, thank you to my family. My parents, my sister, my aunts and uncles, grandparents, cousins, and everyone else who had a hand in fostering my love for reading and learning at an early age. Thank you for always encouraging me to pursue my dreams, even when they seemed out of reach. Thank you for providing me with a constant source of love and laughter. Thank you for everything.
ABSTRACT

An individual’s genetic makeup plays an important role in determining susceptibility to cognitive aging and transition to dementia such as Alzheimer’s disease (AD). Identifying the specific genetic variants that contribute to cognitive aging and AD may aid in early diagnosis of at-risk patients, as well as identify novel therapeutics targets to treat or prevent development of symptoms. Challenges to identifying these specific genes in human studies include complex genetics, difficulty in controlling environmental factors, and limited access to human brain tissue.

Here, we turned to genetically diverse mice from the BXD genetic reference panel (GRP) to overcome some of the barriers traditionally associated with human studies. Using a forward genetics screen, we first identified and validated the gene heterochromatin protein 1 binding protein 3 (Hp1bp3) as a novel modulator of normal cognitive aging. We then demonstrated that targeted knockdown of Hp1bp3 in the hippocampus by 50-75% was sufficient to induce cognitive deficits and transcriptional changes reminiscent of those observed in aging and AD, namely an increase in inflammatory pathways and decrease in neuronal and synaptically-localized transcripts. We also show Hp1bp3 is a translationally relevant target, as transcriptional changes induced by our targeted knockdown significantly overlapped those observed in the aging human brain. In addition, HP1BP3 itself was decreased in the hippocampus of cognitively impaired aging humans. In summary, our results suggest therapeutics designed to target either Hp1bp3 or its downstream effectors may be useful promoting cognitive longevity.

We next expanded on our findings that the BXDs were variably susceptible to cognitive aging and combined the BXD GRP with a well-established transgenic mouse model of AD harboring 5 familial AD mutations, the 5XFAD model. The resulting panel, which we termed the AD-BXDs, consists of genetically diverse yet isogenic F1 mice that all harbor the same high-risk human AD mutations but who differ across the remainder of their genome. We first showed that genetic variation profoundly modified the impact of human AD mutations on both cognitive and pathological phenotypes. We then validated this complex AD model by demonstrating high degrees of genetic, transcriptomic, and phenotypic overlap with human AD. Genetic mapping was used to identify novel genomic loci that modified susceptibility or resilience to cognitive and pathological symptoms of AD. Finally, we used transcriptome profiling to identify gene networks present in the pre-symptomatic mouse brain that predict cognitive performance at an advanced age. Together, the candidates identified through these analyses highlight new potential drivers of susceptibility or resilience to AD and contribute significantly to our understanding of early, potentially causal disease mechanisms.

In summary, work here highlights the utility of genetically diverse mice to elucidate mechanisms underlying complex human disease, namely cognitive aging and AD. In addition, we developed a novel AD mouse population as an innovative and reproducible resource for the study of mechanisms underlying AD. Data presented here
provides convincing evidence that preclinical models incorporating genetic diversity may better translate to human disease. Due to the reproducible nature of the BXDs and resulting AD-BXD models of human aging and AD as well as develop a better understanding of precise mechanisms underlying disease. Together, these resources may ultimately enable precision medicine approaches across a diverse population. Finally, our experimental design is likely to be broadly applicable to mouse models of human disease that incorporate a dominantly inherited high-risk genotype in the form of a transgene or other genetic perturbation, enhancing the general utility of results reported here.
TABLE OF CONTENTS

CHAPTER 1. INTRODUCTION..................................................................................................1
Aging .......................................................................................................................................1
Alzheimer’s Disease ..............................................................................................................1
  Putative mechanisms underlying Alzheimer’s disease dementia ...................................2
    Amyloid cascade hypothesis.............................................................................................2
    Calcium dysregulation.......................................................................................................4
    Inflammation and immune activation................................................................................5
    Additional considerations..................................................................................................6
  Individual variation in disease onset................................................................................6
  Barriers to identifying effective therapeutics.....................................................................7
Animal Models of Alzheimer’s Disease ..............................................................................8
  5XFAD model .....................................................................................................................8
    Pathology .........................................................................................................................8
    Cognitive deficits .............................................................................................................9
    Deficits in neuronal function ..........................................................................................10
    Additional symptoms ......................................................................................................11
    Sensitivity to background genetics ................................................................................11
  Caveats to traditional models of Alzheimer’s disease .......................................................12
Modeling Genetic Diversity in the Mouse ........................................................................14
  Consideration of genetic diversity in AD studies .............................................................15
  Recombinant inbred strains ..............................................................................................16
    BXD genetic reference panel .........................................................................................18
    Utility for systems genetics ............................................................................................18
  Recombinant inbred backcross .........................................................................................19
Hypothesis and Specific Aims ............................................................................................19
  Specific Aim 1: Evaluate the effect of individual genetic differences on age-related cognitive decline. .........................................................................................................................20
  Specific Aim 2: Evaluate the effect of individual genetic differences on AD-related cognitive decline and neuropathology. ........................................................................................................20
  Specific Aim 3: Identify specific genes involved in modifying susceptibility to age and AD-related cognitive decline through QTL mapping and prioritization analyses .................................................................20

CHAPTER 2. SYSTEMS GENETICS IDENTIFIES HP1BP3 AS A MODIFIER OF NORMAL COGNITIVE AGING ..........................................................21
Introduction .........................................................................................................................21
Methods ...............................................................................................................................22
  Animals ..............................................................................................................................22
  Behavioral assays ............................................................................................................22
    Contextual fear conditioning .........................................................................................22
    T-maze ...........................................................................................................................23
  Calculation of memory index for adult BXD strains ......................................................23
  QTL identification and candidate gene selection ............................................................23
Western blots .........................................................24
Mouse samples .........................................................24
Human samples ..........................................................24
Transcriptome analyses .................................................25
RNA sequencing ..........................................................25
Differential expression analysis .........................................25
Gene set enrichment analyses ............................................25
Design of viral vectors ....................................................26
Intrahippocampal injections ............................................27
Statistical analysis and data availability .....................................27

Results .......................................................................27

Genetic mapping identifies an interval on chromosome 4 associated with memory status at midlife .........................................................27
Prioritization of positional candidates identifies Hp1bp3 as putative regulator of cognitive aging across BXD panel ...........................................29
Genetic correlation analyses suggest important functional roles for Hp1bp3 in cognitive aging .................................................................33
Functional validation in a knock-out mouse model confirms novel role for Hp1bp3 in cognitive function .........................................................33
Hippocampus-specific knockdown of Hp1bp3 impairs cognitive function .........................................................36
Hp1bp3 knockdown alters specific aspects of the hippocampal transcriptome .........................................................40
Hp1bp3 selectively regulates a subset of micro-RNAs .........................................................44
Transcriptome-level changes induced by Hp1bp3 knockdown recapitulate those observed in human aging .........................................................46
HP1BP3 is associated with memory function in human aging .........................................................48

Discussion ..................................................................48

Identification and validation of Hp1bp3 as a modifier of cognitive decline .........................................................48
Hp1bp3 effect on cognition is robust to genetic context .........................................................50
Hp1bp3 as a regulator of miRNA biogenesis .........................................................51
Hp1bp3 knockdown recapitulates symptoms of aging and Alzheimer’s disease .........................................................52
Hp1bp3 may play conserved role in humans .........................................................52

CHAPTER 3. HARNESSING GENETIC COMPLEXITY TO ENHANCE MOUSE MODELS OF AD: A PATH TOWARD PRECISION MEDICINE ........54

Introduction ...................................................................54
Methods .......................................................................56
Animals .......................................................................56
Behavioral assays .........................................................57
Y-maze .......................................................................57
Sensorimotor battery .........................................................57
Elevated plus maze .........................................................57
Contextual fear conditioning .........................................................58
Heritability estimates .........................................................58
Enzyme-linked immunosorbent assay (ELISA) .........................................................59
Immunohistochemistry and plaque quantification .........................................................59
Transcriptome analyses .........................................................60
CHAPTER 4. IDENTIFICATION OF GENES AND NETWORKS THAT MODIFY SUSCEPTIBILITY TO ALZHEIMER’S DISEASE

INTRODUCTION.................................................................81

METHODS .........................................................................82

Animals ...........................................................................82

Genetic mapping .............................................................83

Mapping of behavioral and pathological traits ..................83

Mapping of transcriptional traits ......................................83

TRANSCRIPTOME ANALYSES ............................................83

RNA sequencing .............................................................83

Gene set enrichment analysis ..........................................84

Cell-type specific enrichment analysis ..............................84

Weighted gene co-expression network analysis ................84

RESULTS ..........................................................................85

Identification of novel loci modifying AD susceptibility ......85

Inflammation and loss of synaptic genes underlie population-level cognitive decline in AD-BXDs .........................87

Characterization of pre-symptomatic AD transcriptional network .................................................................90

SELECT PRE-SYMPTOMATIC GENE SIGNATURES CORRELATE WITH COGNITIVE FUNCTION LATER IN DISEASE .................................................................................90

Identification of specific drivers of module-trait associations .................................................................93

Genomic region on chromosome two underlies variation in resilience networks ..............................................94

Discussion ......................................................................98

Utility of the AD-BXDs for studying resilience ..................98

The role of modifiers genes in normal aging and AD ........98

Hub genes as candidates for promoting resilience ..............99

Identification of genetic mediators of resilience .................100
CHAPTER 5. CONCLUSIONS AND FUTURE DIRECTIONS .............................................101
Utility of Genetic Diversity in Mouse Preclinical Studies ........................................101
Hp1bp3 as a Mediator of Normal Cognitive Aging ..................................................102
Limitations of the Current Study ...........................................................................102
Genetically Diverse Mouse Models as a Resource for Precision Medicine ..........103
Future Directions ....................................................................................................104

LIST OF REFERENCES ...............................................................................................105

APPENDIX A. SUPPLEMENTARY DATA FOR CHAPTER 2 .................................132
APPENDIX B. SUPPLEMENTARY DATA FOR CHAPTER 3 .................................137
VITA ..........................................................................................................................145
LIST OF TABLES

Table 2-1. Locally (cis) regulated positional candidates.........................................................30

Table 2-2. Demographics for elderly humans used for postmortem hippocampal tissue analysis..........................................................................................................................49

Table 3-1. Heritability estimates for phenotypic traits in AD- and Ntg-BXD. ..........66

Table 3-2. Genes known to confer risk for human AD and their odds ratio across the AD-BXDs.................................................................72

Table 4-1. List of QTLs reaching genome-wide significance across the AD-BXD and Ntg-BXD populations.........................................................................................86

Table A-1. Upstream regulators of observed mRNA changes as predicted by Ingenuity Pathway Analysis.................................................................134
**LIST OF FIGURES**

| Figure 1-1. | Proteolytic processing of APP. | 3 |
| Figure 1-2. | Strain background-dependent variation in amyloid production. | 13 |
| Figure 1-3. | Construction of a set of RI strains from B6 and D2 progenitors. | 17 |
| Figure 2-1. | Identification of *Hp1bp3* as top candidate modulator of cognitive aging. | 28 |
| Figure 2-2. | Effect of *Hp1bp3* genotype on CFM status is age-dependent. | 31 |
| Figure 2-3. | Differences in CFM were not attributable to baseline freezing or post-shock reactivity. | 32 |
| Figure 2-4. | Gene Set Enrichment Analysis (GSEA) suggest important functional roles for *Hp1bp3* in cognitive aging. | 34 |
| Figure 2-5. | *Hp1bp3* KO mice exhibit impairment on hippocampus-dependent long-term and working memory tasks. | 37 |
| Figure 2-6. | Hippocampus-specific knockdown of *Hp1bp3* impairs cognitive function. | 38 |
| Figure 2-7. | Knockdown of *Hp1bp3* alters specific aspects of the hippocampal transcriptome. | 41 |
| Figure 2-8. | *Hp1bp3* KD negatively regulates genes related to neuronal synaptic plasticity. | 43 |
| Figure 2-9. | *Hp1bp3* selectively regulates the miRNA transcriptome *in vivo*. | 45 |
| Figure 2-10. | Transcriptome-level changes induced by *Hp1bp3* knockdown recapitulate those observed in human aging. | 47 |
| Figure 2-11. | HP1BP3 protein is correlated with memory status in aging humans. | 49 |
| Figure 3-1. | Genetic background modifies AD symptoms in a novel transgenic reference panel. | 64 |
| Figure 3-2. | AD-BXD panel is sensitive to variation in known AD risk gene *ApoE*. | 68 |
| Figure 3-3. | Genetic risk score calculated from genotype at known AD risk genes predicts cognitive decline. | 70 |
| Figure 3-4. | Genetic background modifies AD-associated transcriptome. | 74 |
Figure 3-5. Aged AD-BXD transcriptome shows high concordance with late-onset human AD signature.................................................................76

Figure 4-1. Cognitive function declines with age across genetically diverse AD mice. ..........................................................................................88

Figure 4-2. Inflammation and loss of synapses likely underlie population-level cognitive decline in AD-BXDs. ....................................................89

Figure 4-3. Identification of pre-symptomatic gene signatures predictive of AD resilience...............................................................................91

Figure 4-4. Identification of putative target genes that may be targeted to promote resilience networks..........................................................95

Figure 4-5. Identification of genomic regions underlying variation in resilience networks. ................................................................................97

Figure A-1. Delivery of Hp1bp3-shRNA to the hippocampus does not affect cortical or cerebellar levels of HP1BP3........................................132

Figure A-2. KD does not have non-specific effects on anxiety, activity, or body weight.............................................................................133

Figure B-1. Cognitive traits are not directly associated with a variety of peripheral traits...........................................................................137

Figure B-2. AD-BXD mice develop plaques throughout the hippocampus and cortex.138

Figure B-3. Cognitive traits are not directly associated with Aβ42 levels.............139

Figure B-4. Background strain does not significantly modify 5XFAD transgene expression or endogenous App/Psen1 levels.........................140

Figure B-5. Apoe genotype effects CFM, but not CFA, in Ntg-BXD mice. ............142

Figure B-6. AD genetic risk score defined in Ntg-BXD mice shows no relationship with cognitive outcomes......................................................143

Figure B-7. AD-BXD panel exhibits high transcriptional concordance with human AD.................................................................144
### LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5XFAD</td>
<td>Transgenic model of Alzheimer’s disease incorporating 5 familial mutations in the genes APP and PS1.</td>
</tr>
<tr>
<td>AAO</td>
<td>Age at onset</td>
</tr>
<tr>
<td>AAV</td>
<td>Adeno-associated virus</td>
</tr>
<tr>
<td>Aβ</td>
<td>Amyloid-beta</td>
</tr>
<tr>
<td>Aβ40</td>
<td>Amyloid-beta of specifically 40 amino acids in length</td>
</tr>
<tr>
<td>Aβ42</td>
<td>Amyloid-beta of specifically 42 amino acids in length</td>
</tr>
<tr>
<td>ABCA7</td>
<td>ATP-binding cassette transporter A7</td>
</tr>
<tr>
<td>AD</td>
<td>Alzheimer’s disease</td>
</tr>
<tr>
<td>AHP</td>
<td>After-hyperpolarization</td>
</tr>
<tr>
<td>APOE</td>
<td>Apolipoprotein E</td>
</tr>
<tr>
<td>APP</td>
<td>Amyloid precursor protein</td>
</tr>
<tr>
<td>B6</td>
<td>C57BL/6J mouse strain</td>
</tr>
<tr>
<td>BDNF</td>
<td>Brain derived neurotrophic factor</td>
</tr>
<tr>
<td>BXD</td>
<td>Mouse genetic reference panel derived from B6 and D2 progenitors</td>
</tr>
<tr>
<td>CA1</td>
<td>Cornu Ammonis area 1 of the hippocampus</td>
</tr>
<tr>
<td>CFA</td>
<td>Contextual fear acquisition</td>
</tr>
<tr>
<td>CFC</td>
<td>Contextual fear conditioning</td>
</tr>
<tr>
<td>CFM</td>
<td>Contextual fear memory</td>
</tr>
<tr>
<td>D2</td>
<td>DBA/2J mouse strain</td>
</tr>
<tr>
<td>ELISA</td>
<td>Enzyme-linked immunosorbent assay</td>
</tr>
<tr>
<td>EPM</td>
<td>Elevated plus maze</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>ES</td>
<td>Enrichment score</td>
</tr>
<tr>
<td>F1</td>
<td>First filial generation</td>
</tr>
<tr>
<td>F2</td>
<td>Second filial generation</td>
</tr>
<tr>
<td>FAD</td>
<td>Familial Alzheimer’s disease</td>
</tr>
<tr>
<td>FDR</td>
<td>False discovery rate</td>
</tr>
<tr>
<td>GEO</td>
<td>Gene Expression Omnibus</td>
</tr>
<tr>
<td>GO</td>
<td>Gene Ontology</td>
</tr>
<tr>
<td>GRP</td>
<td>Genetic reference panel</td>
</tr>
<tr>
<td>GRS</td>
<td>Genetic risk score</td>
</tr>
<tr>
<td>GSEA</td>
<td>Gene set enrichment analysis</td>
</tr>
<tr>
<td>GWAS</td>
<td>Genome-wide association study (or studies)</td>
</tr>
<tr>
<td><em>Hp1bp3</em></td>
<td>Heterochromatin protein 1 binding protein 3</td>
</tr>
<tr>
<td>IHC</td>
<td>Immunohistochemistry</td>
</tr>
<tr>
<td>IPA</td>
<td>Ingenuity Pathway Analysis</td>
</tr>
<tr>
<td>KD</td>
<td>Knockdown</td>
</tr>
<tr>
<td>KEGG</td>
<td>Kyoto Encyclopedia of Genes and Genomes</td>
</tr>
<tr>
<td>KO</td>
<td>Knock-out</td>
</tr>
<tr>
<td>LOAD</td>
<td>Late-onset Alzheimer’s disease</td>
</tr>
<tr>
<td>LTD</td>
<td>Long-term depression</td>
</tr>
<tr>
<td>LTP</td>
<td>Long-term potentiation</td>
</tr>
<tr>
<td>MAPT</td>
<td>Microtubule-associated protein tau</td>
</tr>
<tr>
<td>ME</td>
<td>Module eigengene</td>
</tr>
<tr>
<td>MMSE</td>
<td>Mini-mental status exam</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------------------------------------------</td>
</tr>
<tr>
<td>MRI</td>
<td>Magnetic resonance imaging</td>
</tr>
<tr>
<td>MSigDb</td>
<td>Molecular Signatures Database</td>
</tr>
<tr>
<td>NES</td>
<td>Normalized enrichment score</td>
</tr>
<tr>
<td>NOR</td>
<td>Novel object recognition</td>
</tr>
<tr>
<td>Ntg</td>
<td>Non-transgenic</td>
</tr>
<tr>
<td>PET</td>
<td>Positron emission tomography</td>
</tr>
<tr>
<td>PMI</td>
<td>Post-mortem interval</td>
</tr>
<tr>
<td>PSEN1</td>
<td>Presenilin1</td>
</tr>
<tr>
<td>PSEN2</td>
<td>Presenilin2</td>
</tr>
<tr>
<td>QTL</td>
<td>Quantitative trait locus</td>
</tr>
<tr>
<td>RI</td>
<td>Recombinant inbred</td>
</tr>
<tr>
<td>RIB</td>
<td>Recombinant inbred backcross</td>
</tr>
<tr>
<td>SNP</td>
<td>Single nucleotide polymorphism</td>
</tr>
<tr>
<td>TPM</td>
<td>Transcripts per million</td>
</tr>
<tr>
<td>WGCNA</td>
<td>Weighted gene co-expression network analysis</td>
</tr>
<tr>
<td>WT</td>
<td>Wild-type</td>
</tr>
</tbody>
</table>
CHAPTER 1. INTRODUCTION

Aging

With the advent of better medical care and other societal advances, the number of elderly living in the United States is expected to drastically increase in the coming years. Specifically, the number adults over the age of 65 is expected to more than double, from 40.2 million in 2010 to 88.5 million in 2050 (Vincent, 2010). As aging is the leading risk factor for a number of diseases, particularly age-related dementias such as Alzheimer’s disease (AD), there is a critical public health need for a better understanding of mechanisms that drive healthy brain aging versus those that promote a transition to pathological aging.

The aging process itself has been associated with a number of structural and physiological changes in the brain, including less coordinated activation of memory-relevant brain regions, decreased numbers of dendritic spines, and altered synaptic transmission (Liu et al., 1996; Yankner et al., 2008). These changes are often accompanied by some degree of impairments in cognitive function (Bishop et al., 2010; Petersen et al., 1992). These impairments can include deficits in working memory, short-term recall, processing speed, and spatial memory (Yankner et al., 2008). While in the context of normal aging these impairments generally do not interfere with daily living, the degree of cognitive impairment that occurs with age is highly variable across the population. For example, there have been reports of so-called “super-agers” that progress to extreme age with no indication of cognitive decline (Harrison et al., 2012), while others exhibit more noticeable decline and eventual descent into mild cognitive impairment or dementia. Heritability studies indicate that both cognitive function in old age (McClearn et al., 1997; McGue and Christensen, 2001) and risk of developing dementia (Gatz et al., 1997; Hinz and Geschwind, 2017) are strongly influenced by genetics. Thus, studies aimed at identifying specific genetic factors associated with cognitive decline are likely to provide effective therapeutic targets that can be manipulated to promote cognitive longevity.

Alzheimer’s Disease

The risk of developing AD increases exponentially with age, with as many as half of individuals over the age of 85 being afflicted with this devastating disease (Rocca et al., 2011). By 2050, the number of people with AD is projected to increase nearly 30% to 13.8 million, costing the United States alone an estimated $1.1 trillion in associated healthcare costs (Alzheimer’s Disease Facts and Figures, 2018). Thus, there is a critical need to better understand disease causes, underlying risk factors, and mechanisms driving disease progression. AD is a neurodegenerative disease characterized by a progressive loss of cognitive function and two pathological hallmarks, extracellular amyloid plaques and intracellular neurofibrillary tangles made up of hyperphosphorylated tau protein (Selkoe, 1991). In contrast to normal aging, AD patients exhibit an extensive degree of
neuron loss, particularly in vulnerable areas such as the hippocampus (Moodley and Chan, 2014; Terry et al., 1991). In addition, the degree and rate of cognitive impairment is much greater in AD patients than in those experiencing normal age-related cognitive decline, extending to the domains of episodic memory and executive function (Deary et al., 2009) and beginning to interfere with daily living abilities.

Although age is the greatest risk factor for developing AD, it is increasingly clear that genetics and family history play a large role. The heritability of AD is estimated to be in the range of 50-80%, indicating an individual’s susceptibility or resilience to disease is, at least in part, determined by heritable DNA variants (Gatz et al., 1997). There are generally considered to be two types of AD, the first being the autosomal dominant form of the disease known as familial AD (FAD), now known to be caused by inherited mutations in the genes encoding for amyloid precursor protein (APP) (Goate et al., 1991) or presenilin proteins 1 or 2 (PSEN1 and PSEN2, respectively, Tanzi et al., 1996). FAD typically presents early in life, before the age of 65 and as such, is also known as early-onset AD. In contrast, the second form of the disease is known as late-onset AD (LOAD), as disease onset typically occurs later in life after the age of 65. While a number of risk factors have been identified for LOAD, including specific gene mutations (Lambert et al., 2013) and environmental factors (Livingston et al., 2017), precise causes are still unknown. However, decades of research have provided insight into a variety of mechanisms that may promote the transition from healthy normal aging to pathological AD.

Putative mechanisms underlying Alzheimer’s disease dementia

Amyloid cascade hypothesis

Accumulation of extracellular plaques is one of the pathological hallmarks of AD (Selkoe, 1991). The primary component of these plaques is a peptide 42 amino acids in length known as beta-amyloid 42 (Aβ42). Aβ42 is produced by the processing of APP, a transmembrane protein which can be processed in one of two ways, either down a route leading to the production of beta-amyloid (referred to as the amyloidogenic route) or a route that produces alternative products that are not aggregation-prone, generally referred to as the non-amyloidogenic route (Figure 1-1, Haass et al., 2012). Mutations in APP that have been associated with FAD often occur within or around the domain that eventually becomes the Aβ peptide. These mutations have a variety of effects, including 1) increasing the production of total Aβ, as is the case with the Swedish mutation (Mullan et al., 1992), 2) specifically increasing the production of Aβ42, as with the Florida (Eckman et al., 1997) and London (Goate et al., 1991) mutations, or 3) increasing the aggregation propensity of Aβ, as with the Arctic mutation (Nilsberth et al., 2001). The two other genes associated with FAD, PSEN1 and PSEN2, are components of the gamma-secretase complex, which mediates the second step in APP cleavage (Figure 1-1). Cleavage of APP by gamma-secretase is not specific, and can occur at a number of different amino acids along the protein, likely through a multi-step process (Haass et al., 2012). FAD
Figure 1-1.  Proteolytic processing of APP.
Processing of APP can proceed down an (left) anti-amyloidogenic route or (right) an amyloidogenic route. The amyloidogenic route produces toxic amyloid species, particularly species 42 amino acids in length (Aβ42).
mutations in PSEN1 and PSEN2 generally increase the relative amount of toxic Aβ42 relative to the shorter and less aggregation-prone amyloid species containing only 40 amino acids (Aβ40, De Strooper et al., 2012).

As mutations identified as disease-causative in FAD have all been observed to drive the over-production of Aβ42, a leading hypothesis in the field is that amyloid accumulation is a causal step in a cascade of events eventually leading to the synaptic dysfunction, cognitive impairment, and neurodegeneration observed later in disease (Hardy and Higgins, 1992). Besides the age at onset, FAD clinically presents very similarly to the more common LOAD (Hall and Roberson, 2012; LaFerla and Green, 2012). As such, the amyloid hypothesis has been generalized to apply to all forms of the disease and Aβ has long been assumed to be the causal molecule in AD (Reitz, 2012). In support of this hypothesis, both in vitro and in vivo studies have shown that Aβ42 is neurotoxic and causes eventual cell death (Harkany et al., 2000; Kowall et al., 1992). In addition, overexpression of mutated human APP in mouse and flies produces cognitive deficits and neurodegeneration (Greeve et al., 2004; Oakley et al., 2006), suggesting Aβ accumulation is sufficient to produce symptoms reminiscent of the human disease. However, the correlation between the most problematic AD symptom, cognitive impairment, and Aβ is relatively weak in humans (Villemagne et al., 2011). Indeed, a high percentage of elderly individuals that show no cognitive decline, even in very old ages, have been shown to have pathological levels of amyloid in their brains, suggesting that amyloid alone is not always sufficient to drive the progression and incidence of AD (Herrup, 2015). Perhaps most clinically relevant, none of the clinical trials targeting Aβ in AD have been successful in halting or slowing disease progression (van Dyck, 2018). As such, it is still unclear whether Aβ itself is the initiating step driving AD onset, or simply a consequence of earlier disease-causative mechanisms.

**Calcium dysregulation**

In addition to the pathological accumulation of plaques and tangles in Alzheimer’s disease, dysregulation of intracellular calcium (Ca²⁺) homeostasis has been widely observed in both sporadic and familial AD (reviewed in LaFerla, 2002). Notably, alterations in Ca²⁺ signaling were detected in fibroblasts of individuals carrying FAD mutations prior to symptom onset, suggesting Ca²⁺ dysregulation plays a causal role in disease pathogenesis (Etcheberriagaray et al., 1998). In further support of a causal role for Ca²⁺ dyshomeostasis, each of the genes associated with FAD, in addition to driving increased production of Aβ42, have also been shown to independently alter Ca²⁺ signaling (LaFerla, 2002), particularly in the endoplasmic reticulum and in maintenance of intracellular Ca²⁺ levels (Bezprozvanny and Mattson, 2008). This “calcium hypothesis” of AD (Khachaturian, 1989) also helps explain why aging exponentially increases the risk for AD; it is thought that as neurons age, their ability to maintain regulation of Ca²⁺ levels decreases due to inefficient energy metabolism and accumulating oxidative stress (Supnet and Bezprozvanny, 2010).

As calcium is utilized by neurons to control a wide variety of functions, including membrane excitability, neurotransmitter release, gene expression, and cellular growth,
Ca\textsuperscript{2+} dysregulation likely leads to a variety of altered cellular processes (Berridge, 1998; Supnet and Bezprozvanny, 2010). In particular, Ca\textsuperscript{2+-}dependent signaling pathways play essential roles in neuronal excitability and plasticity, the leading mechanisms hypothesized to underlie learning and memory in the brain (Kandel, 2004). Specifically, increased intracellular Ca\textsuperscript{2+} in aged and AD neurons, particularly those from rodent models (Kaczorowski et al., 2011; Kumar and Foster, 2004), has been shown to result in an increased after-hyperpolarization (AHP) period following action potential firing. An increased AHP is linked to decreased intrinsic neuronal excitability, and thus, and an increased threshold for the induction of long-term potentiation (LTP) (Zhang and Linden, 2003). LTP is a process by which the connection of neurons is strengthened following coordinated activity and can be induced by brief, large magnitude rises in post-synaptic Ca\textsuperscript{2+} (Foster, 2007). The opposing process, known as long-term depression (LTD), weakens the connections between neurons and can be trigged by a modest and prolonged rise in intracellular Ca\textsuperscript{2+} (Foster, 2007). As such, changes in Ca\textsuperscript{2+} homeostasis that occur in the aging and diseased brain can have profound implications for neuronal communication and synaptic function. As synaptic dysfunction is intimately linked with cognitive decline in AD (Selkoe, 2002), these findings highlight Ca\textsuperscript{2+} dysfunction as a potential early causative driver of clinical AD symptoms.

**Inflammation and immune activation**

Neuroinflammation has long been observed in AD (McGeer et al., 1989). However, it is only relatively recently that genetic studies have implicated the immune system as potentially playing a causal role in the development of AD (Selkoe and Hardy, 2016). A number of genome-wide association studies (GWAS) attempting to link gene variants with susceptibility for AD have identified hits in, or around, genes with established functions in the immune system. These include triggering receptor expressed on myeloid cells 2 (TREM2), CD33, and complement receptor 1 (CR1), among others (Bertram et al., 2008; Lambert et al., 2009; Lambert et al., 2013). Indeed, statistically significant enrichments for immune-related genes have been observed among AD-related GWAS hits (Jones et al., 2010), suggesting variation in the innate immune response is critical for determining individual susceptibility to AD development and progression.

Several hypotheses exist as to how the immune system may modulate susceptibility to AD. Microglia, the resident immune cell of the brain, have been shown to phagocytose amyloid deposits early in disease (Jevtic et al., 2017), and their activation is initially thought to be beneficial. However, microglia are also known to release pro-inflammatory cytokines when activated (Orihuela et al., 2016). As such, prolonged activation in response to cellular debris such as amyloid may lead to sustained neuroinflammation, resulting in cellular stress, reactive oxygen species, and neuronal damage (Heneka et al., 2014; Jevtic et al., 2017). GWAS hits in key microglia genes including CD33 and TREM2 implicate variation in these response pathways as critical for disease pathogenesis. In addition, the complement system has emerged as another part of the innate immune system that may critically modulate AD progression. The complement system spans multiple cell types and involves signaling molecules produced by both neurons and astrocytes, as well as a number of receptors expressed on microglia (Jevtic et
al., 2017). Typically, this system is induced to facilitate synaptic pruning and neural circuit maintenance (Bahrini et al., 2015; Schafer et al., 2012), but current data suggests that complement becomes over-activated in AD and leads to aberrant synaptic pruning and eventual cell death (Jevtic et al., 2017). Additional evidence has suggested a role for peripheral immune cells, including mononuclear phagocytes, neutrophils, and lymphocytes, either through true peripheral responses or infiltration into the central nervous system (Jevtic et al., 2017). While much remains to be elucidated about the role of the immune system in AD, there exists a growing appreciation for the role of a multitude of cell types in the complex etiology of AD.

**Additional considerations**

In addition to the putative mechanisms discussed above, a number of additional processes have been hypothesized as critical to the development of AD. Much of this insight comes from the identification of genetic variants linked to AD risk (Karch and Goate, 2015). For example, apolipoprotein E (APOE), particularly the e4 allele, is the strongest known risk factor for LOAD, and its central role in cholesterol metabolism highlights the potential importance for this pathway in mediating AD risk (Corder et al., 1993; Karch and Goate, 2015). Several additional risk genes including clusterin (CLU) and ATP-binding cassette transporter A7 (ABCA7) have important roles in cholesterol metabolism, including in the regulation of lipid transport out of cells and into lipoprotein particles (Jones and Jomary, 2002; Kim et al., 2008). Genes with essential roles in endocytosis and synaptic function have also been implicated in mediating risk for AD, including bridging integrator 1 (BIN1), phosphatidylinositol binding clathrin assembly protein (PICALM), and CD2-associated protein (CD2AP) (Karch and Goate, 2015; Lambert et al., 2013). Together, these GWAS hits highlight intracellular trafficking (Pant et al., 2009), vesicle recycling (Di Paolo et al., 2002), and cytoskeletal reorganization (Dustin et al., 1998) as processes that may be targeted to protect against AD. As the size of GWAS continue to increase, it is likely additional genes and pathways will be implicated in AD pathogenesis, providing new therapeutic opportunities and insight into the genetic and molecular causes of disease.

**Individual variation in disease onset**

Although the precise cause(s) of AD are not well understood, it has become increasingly clear that AD is a highly complex disease, largely driven by a combination of a number of small effect gene variants and various environmental factors (Livingston et al., 2017). In addition, AD is highly heterogeneous in presentation and progression, with variations observed in patterns of neuronal atrophy, amyloid deposition, and cognitive impairments (Lam et al., 2013). Even among FAD patients carrying what are typically thought to be highly penetrant mutations in APP and PSEN1, the age at which patients first show disease onset can range from 20 to 70 years of age (Ryman et al., 2014). Moreover, as mentioned previously, elderly individuals can carry pathological loads of amyloid protein while remaining cognitively intact. These resilient individuals, who exhibit better than expected cognitive function based on the presence of
neuropathology and/or high-risk gene mutations, represent a clinically interesting subset of the population (Hohman et al., 2017; Hohman et al., 2016). Understanding the genetic and molecular causes of resilience may help identify new therapeutic opportunities and enable understanding of how we may delay or prevent disease onset in a broader patient population.

**Barriers to identifying effective therapeutics**

Studies focused on individuals resilient to developing cognitive symptoms related to extreme age or AD pathology represent prime opportunities to identify and understand factors that may protect against age-related cognitive decline and dementia. However, a number of barriers associated with human studies have prohibited a detailed understanding of the mechanisms that underlie resilience. In particular, resilient individuals who remain cognitively intact (despite genetic risk for AD) rarely enter the clinic in the absence of symptoms. If they do enter the clinic, they are likely to be enrolled in clinical trials or observational studies as normal controls due to a lack of cognitive symptoms, potentially confounding genetic association studies. In addition, it is difficult to say which individuals exhibit incomplete resilience and will eventually go on to develop AD symptoms (Moulder et al., 2013). Even among FAD patients with later-than-predicted disease onset (Ryman et al., 2014) and who are monitored across the lifespan due to family history in cohorts such as the Dominantly Inherited Alzheimer’s Network (Moulder et al., 2013), sample sizes are general small and ill-powered for genome-wide discovery of genes that modify disease trajectory.

Across both resilient and susceptible individuals, there is a general lack of access to memory-relevant brain tissue at early disease time points, which precludes the identification of molecular mechanisms driving the onset of disease. While a number of studies have attempted to utilize biomarkers from peripheral tissue such as blood, plasma, and cerebrospinal fluid, how these markers relate to mechanisms in the brain is still poorly understood (Lashley et al., 2018; Shi et al., 2018). Instead, a large majority of studies rely on postmortem tissue from late-stage disease (Perl, 2010). As such, what is represented in these samples is often consequential signatures resulting from decades of disease, rather than early causative mechanisms that may be targeted, manipulated, or halted to stop disease progression before irreparable damage is done.

In addition, the human genome is incredibly complex, which complicates the execution and interpretation of GWAS aimed at identifying specific variants associated with disease (Lowe and Reddy, 2015). Specifically, a number of variants associated with AD risk fall in non-coding regions of the genome and/or areas of strong linkage disequilibrium (Karch et al., 2014), making assigning causality to any given gene or regulatory element in the region difficult. Moreover, due to the highly polymorphic nature of the human genome, GWAS suffer a high multiple testing correction burden (Johnson et al., 2010) and as such, require large numbers of cases and controls to yield significant and meaningful results (Lambert et al., 2013). Finally, a variety of environmental factors are known to profoundly affect risk for cognitive decline and
dementia, including physical factors such as diet, inactivity, smoking, diabetes and obesity as well as additional factors such as educational attainment, socioeconomic status, social isolation, and depression (Livingston et al., 2017). Lifetime exposure to, and experience with, these environmental risk factors varies profoundly across individuals and unfortunately often goes unreported or unaccounted for in genetic studies.

**Animal Models of Alzheimer’s Disease**

Due to many of the barriers associated with human studies, animal models, particularly mouse models, have emerged as powerful way to study the pathophysiology of AD (LaFerla and Green, 2012). Many of the available mouse models take advantage of *APP* and *PSEN1* mutations known to cause FAD and incorporate transgenes encoding various human mutations, either in isolation or in combination, to recapitulate aspects of the human disease. Several models also incorporate human mutations in the microtubule-associated protein tau (*MAPT*, Platt et al., 2013). Although *MAPT* mutations have not been causally linked to AD in human populations, their inclusion aids in driving the pathological accumulation of hyperphosphorylated tau, and ultimately, neurofibrillary tangles required for AD diagnosis (Jankowsky and Zheng, 2017). While no animal model fully recapitulates all aspects of the human condition, together they have led to a number of advances in our understanding of mechanisms underlying specific aspects of disease, including pathological plaque and tangle formation (LaFerla and Green, 2012), neurodegeneration (Grandi et al., 2018), and cognitive dysfunction (Webster et al., 2014).

**5XFAD model**

One of the most well-characterized models of AD is the 5XFAD model, originally developed in the laboratory of Robert Vassar at Northwestern University (Oakley et al., 2006) and named due to its inclusion of five mutations associated with human FAD. Specifically, three mutations in *APP* (Swedish; K670N/M671L, Florida; I716V, and London; V717I) and two in *PSEN1* (M146L and L286V) were introduced into *APP* and *PSEN1* cDNAs by site-directed mutagenesis, and then cloned into exon 2 of the mouse *Thy1* transgene cassette to drive overexpression of mutated *APP* and *PSEN1* in the mouse brain (Oakley et al., 2006). The transgenes were then injected into hybrid B6SJL mouse embryos (Oakley et al., 2006) and co-integrated at a single insertion site now known to be located on mouse chromosome three (Goodwin, 2017). The 5XFAD model is known to recapitulate many aspects of the human disease and was recently shown to more closely resemble human AD on a transcriptional level than many currently available alternative models (Hargis and Blalock, 2017).

**Pathology**

On the original B6SJL background, Aβ42 was detectable in the brains of 5XFAD animals by two months of age. Deposition begins intraneuronally and ultimately progresses to plaque formation, first in layer five of the cortex and in the subiculum
region of the hippocampus, ultimately spreading throughout the brain (Oakley et al., 2006). Synaptic markers including synaptophysin, syntaxin, and postsynaptic density-95 were decreased in accordance with marked neuronal loss by nine months of age (Eimer and Vassar, 2013; Oakley et al., 2006). The observed neuron loss in 5XFAD mice is particularly notable, as many animal models of AD do not exhibit this hallmark symptom of disease (Hall and Roberson, 2012). Specifically, neurodegeneration occurs primarily in layer 5 of the cortex and the subiculum, particularly in large pyramidal neurons that also show evidence of intraneuronal accumulation of Aβ42 (Eimer and Vassar, 2013). Data suggests neurons die primarily by an apoptotic mechanism, as indicated by elevations of caspase-3 (Eimer and Vassar, 2013), although the exact mechanism(s) driving neuronal death are not well established.

In addition to Aβ42 accumulation and neuron loss, 5XFAD mice also exhibit extensive gliosis, as indicated by increased densities of both astrocytes and microglia in regions surrounding plaque deposits (Oakley et al., 2006). Transcriptional analyses using RNA from the cortex and hippocampus of female 5XFAD mice highlight an increase in inflammatory and immune markers as predominant feature of 5XFAD brains by four months of age (Landel et al., 2014). Spatially, the pattern of gliosis is observed to closely follow the distribution of amyloid deposits (Oakley et al., 2006), although some data suggest intrinsic changes occur in microglia prior to the onset of amyloid deposition (Boza-Serrano et al., 2018). Overall, 5XFAD mice capture a number of aspects of AD neuropathology, including elevated Aβ42, neuron loss, synaptic degeneration, and gliosis, suggesting this model provides a useful foundation to begin to understand factors driving the onset of these pathologies and how these aspects of disease influence cognitive decline and dysfunction (Oakley et al., 2006). Of note, tangle pathology is not generally observed in 5XFAD mice, although hyperphosphorylation of tau at certain sites has been reported to occur (Kanno et al., 2014), and in some studies, the 5XFAD model has been crossed to tau transgenics in order to generate a more complete model of the human disease (Herard et al., 2014)

**Cognitive deficits**

5XFAD mice display hallmark signs of cognitive dysfunction typically associated with human AD. 5XFAD mice first show deficits on spatial working memory tasks such as the y-maze spontaneous alternation task (Oakley et al., 2006). Working memory is typically assessed using verbal recall assays in human patients, but alternation tasks in rodents instead take advantage of a rodent’s natural tendency to explore novel environments in search of food, water, or shelter. An animal’s ability to choose a new arm over an arm that has been previously explored is measured as an index of spatial working memory (Lalonde, 2002; Webster et al., 2014). Deficits in this domain appear in 5XFAD mice as early as four months of age (Oakley et al., 2006) and are maintained out to 12-14 months of age (Webster et al., 2014). Spatial working memory in the 5XFAD model has also been evaluated using the Morris Water Maze, in which mice are placed in a pool of water and the time to first learn the location of an elevated escape platform is measured, as well as the time spent in the area of the escape platform after its removal during a probe trial. Relative to their non-transgenic littermates, 5XFAD mice take longer
to learn the location of the escape platform and exhibit decreased memory for escape platform location during testing trials by 6-9 months of age (Schneider et al., 2014; Webster et al., 2014).

Cognitive deficits in the 5XFAD model progress with age, extending to cognitive domains of associative and recognition memory by 5-6 months of age (Webster et al., 2014). Associative memory in the 5XFAD model has been evaluated extensively using contextual fear conditioning (CFC) paradigms in which a mouse is placed in a training chamber and receives a series of mild foot shocks, either alone or paired with a preceding stimulus such as light or sound. The animal learns to associate the training chamber and/or the conditioned stimulus with the aversive experience of receiving foot shocks. The degree of associative learning and memory can be measured by the amount of time spent exhibiting a species-specific defensive behavior known as freezing (Curzon et al., 2009; Fanselow, 1980; Webster et al., 2014) during the training and testing trials, respectively. 5XFAD mice have been shown to exhibit reduced freezing relative to their non-transgenic littermates around 6-8 months of age, indicating impairments in contextual fear memory (CFM), consolidation, and/or recall (Kaczorowski et al., 2011; Ohno et al., 2006).

In addition, 5XFAD mice have been reported to exhibit later deficits in recognition memory as measured by performance on the novel object recognition (NOR) task by 9-11 months of age (Tohda et al., 2012; Webster et al., 2014). Similar to the alternation tests of working memory, the NOR task takes advantage of a mouse’s tendency to explore novel environments and objects. Specifically, in the NOR task, a mouse is exposed to two objects and allowed to familiarize itself with each object, typically spending equal time exploring both. During the testing trial, typically following a delay period of anywhere from 1-24 hours, one of the now-familiar objects is replaced with a new object. A mouse with intact recognition memory will recognize the “familiar” object and spend more time exploring the “novel” object. Aged 5XFAD mice are unable to discriminate familiar objects from novel objects and do not display a preference for the novel object as expected from a mouse with intact recognition memory (Tohda et al., 2012; Webster et al., 2014). As human patients with AD exhibit deficits in both fear conditioning (Hamann et al., 2002) and recognition memory (Didic et al., 2013; Ferrucci et al., 2008), both fear conditioning and NOR in 5XFAD mice may represent translationally relevant behavioral assays.

**Deficits in neuronal function**

A number of neurophysiological deficits have been observed in the 5XFAD model, providing a putative explanation for the variety of cognitive impairments observed throughout the lifespan. For example, hippocampal neurons from 5XFAD mice have been observed to be intrinsically less excitable than neurons from age-matched non-transgenic (Ntg) littermates (Kaczorowski et al., 2011), as indicated by an enhanced AHP period following action potential firing. This reduction in intrinsic excitability, which was shown to be Ca$^{2+}$-dependent (Kaczorowski et al., 2011), is accompanied by a failure of 5XFAD hippocampal neurons to undergo the expected degree of learning-related
modulation of neuronal excitability. Intrinsically, neurons from 5XFAD mice with cognitive deficits fail to modify their AHP duration and amplitude relative to Ntg littermates and young animals (Kaczorowski and Disterhoft, 2009; Kaczorowski et al., 2011). In terms of synaptic plasticity, both baseline synaptic transmission and LTP is impaired at Schaffer collateral-CA1 synapses from 5XFAD mice (Kimura and Ohno, 2009). Overall, these data suggest that reductions in intrinsic and synaptic plasticity prevent 5XFAD hippocampal neurons from becoming part of the neural network that is recruited during information storage (Neuner et al., 2015).

Interestingly, these impairments are age-dependent, as they are largely have not been observed in mice less than four months of age (Kaczorowski et al., 2011; Kimura and Ohno, 2009). Age-related declines in learning-related brain-derived neurotrophic factor (BDNF)-tyrosine kinase B signaling pathways have also been observed in aged 5XFAD mice, providing some insight into molecular mechanisms that may underlie alterations in neuronal function (Kimura et al., 2010). In addition, alterations in structural connectivity (Jeon et al., 2018) and numbers of dendritic spines (Crowe and Ellis-Davies, 2014) have been observed in 5XFAD mice, suggesting changes in neuronal structure and function may extend beyond biophysical alterations observed by electrophysiology.

Additional symptoms

A number of additional AD-related symptoms have been described in the 5XFAD mouse model, several of which parallel symptoms observed in the human condition (Buchman and Bennett, 2011; Franzblau et al., 2013; Kessler et al., 1991; Wirths and Bayer, 2008). In particular, 5XFAD mice exhibit age-dependent motor deficits, which includes deficits in limb clapping response, ability to cross a narrow balance beam, motor coordination and learning on the rotarod, gait coordination, and swim speed (Jawhar et al., 2012; O’Leary et al., 2018). Interestingly, no deficit in the ability of peripheral motor nerves to drive muscle output was observed (O’Leary et al., 2018), suggesting the deficits result primarily from a failure of the CNS to coordinate motor activity. In addition, brains of aged (13 month-old) 5XFAD mice exhibited significantly impaired glucose uptake measured by uptake of radiolabeled glucose on positron emission tomography (PET) imaging (Macdonald et al., 2014). These changes in blood glucose coincide with reduced expression of glucose transporter 1 (Slc2a1) in 5XFAD brains, particularly at the blood-brain barrier (Ahn et al., 2018). Finally, alterations in the cerebrovascular structure of 5XFAD mice have been extensively observed, including damaged brain blood vessels (Ahn et al., 2018), loss of pericytes (Giannoni et al., 2016), and amyloid deposition in blood vessels (Giannoni et al., 2016). Taken together, these and other data reported in the literature suggest the 5XFAD model recapitulates multiple aspects of human AD, providing valuable opportunities to identify and understand mechanisms that may drive disease susceptibility across species.

Sensitivity to background genetics

Notably, many of the symptoms discussed above were characterized on the B6SJL mixed background strain on which the 5XFAD model was originally developed (Oakley et al., 2006). As genetics is known to play a significant role in determining an
individual’s risk for developing AD symptoms, (Gatz et al., 2006), it is not necessarily surprisingly that a substantial degree of variability is observed within this single mixed background strain. For example, Eimer and Vassar (2013) showed two distinct groups of 5XFAD mice at 6 months; 1) a group that had noticeable neuron loss relative to Ntg littermates and 2) a group that was protected from noticeable neuron loss despite carrying the 5XFAD transgene (Eimer and Vassar, 2013; Oakley et al., 2006). In addition, B6SJL
5XFAD mice can be divided into two separate groups in terms of behavioral performance; 1) those that can successfully learn a CFC task at 8 months (strong-learners), and 2) those that exhibit fear memory deficits at the same age (weak-learners) (Kaczorowski et al., 2011). This behavioral variation can be mapped back to variation in neuronal physiology, as fear deficits in weak-learners were associated with a larger AHP in hippocampal neurons and decreased intrinsic excitability (Kaczorowski et al., 2011). In further support of the idea that the 5XFAD transgene is sensitive to background genetic variants, the onset of Aβ42 plaque deposition is significantly delayed when the B6SJL
5XFAD mice are backcrossed to pure C57BL/6J (B6) mice (Figure 1-2). As specific genetic and molecular mechanisms influencing AD onset and progression remain poorly understood, the identification of precise genetic variants responsible for modifying the expressivity and penetrance of 5XFAD mutations on pathological, behavioral, and physiological symptoms of AD would likely significantly increase our understanding of disease pathophysiology.

Caveats to traditional models of Alzheimer’s disease

Despite the insight into AD provided by animal models, a number of limitations to these resources exist. Perhaps most noticeably, no single mouse model fully recapitulates all aspects of human AD (Drummond and Wisniewski, 2017). For example, the 5XFAD model lacks noticeable tangle pathology, while other models do not exhibit robust neuronal loss. Concern has also been raised about the constructs used to model AD in mice (Drummond and Wisniewski, 2017; Onos et al., 2016). Specifically, transgenes typically incorporate mutations in either APP or PSEN1 that have been identified as causal in FAD families. As FAD represents a small subset of all AD, questions have been raised about the generalizability of findings from mouse models utilizing these mutations. In addition, multiple FAD mutations are often utilized in a single construct in order to accelerate the development of pathology. As FAD mutations are rarely, if ever, observed in combination, this practice further fosters concerns of generalizability. However, multiple studies have suggested that besides the age of onset, FAD and LOAD are quite similar at the clinical and pathological levels (Hall and Roberson, 2012; Laferla and Green, 2012), suggesting insight into disease mechanisms gained from FAD-based animal models may in fact be applicable to various forms of disease.

An additional caveat to using mouse models of AD is the lack of well-defined outcome measures that translate successfully to human populations. As no drug has yet been successful in clinical trials, evaluating the predictive validity of any given outcome
Figure 1-2.  Strain background-dependent variation in amyloid production. Strain background modifies rate at which toxic Aβ42 is produced in the 5XFAD transgenic model of AD. (A) IHC utilizing antibodies targeted against Aβ42 was performed on brains sections from both B6-5XFAD and B6SJL-5XFAD mice (n = 5) at 2, 4, 6, and 8 months of age. Representative images are shown here. (B) The number of Aβ42-positive plaques per square millimeter of tissue was quantified using ImageJ Particle Analysis. By 8 months of age, the two background strains exhibited significantly different numbers of plaques throughout the brain. Modified with permission from publicly available data from the Mutant Mouse Resource and Research Centers. Retrieved from https://www.mmrrc.org/admin/loginCheck.php?doc=34848/34848%2034840%20data.pdf&access=public%27%20https://www.mmrrc.org/catalog/sds.php?mmrrc_id=34848. Accessed 01/19/2019.
has been difficult. Initial studies of AD mouse models focused on pathology, particularly plaque pathology, as a main outcome measure (Hall and Roberson, 2012). However, as Aβ levels do not correlate well with cognitive outcomes in human populations (Arriagada et al., 1992), drugs that exclusively target and clear Aβ pathology in animal models are unlikely to yield valuable therapeutic options in humans. Functional measures such as electrophysiological or behavioral outcomes may produce more relevant results, but these types of outcomes present their own challenges in translating findings to human studies. Electrophysiological data is difficult to obtain in human patients, although imaging technologies including high-resolution functional magnetic resonance imaging (MRI) may provide a suitable estimate of in vivo brain function, and can be used in both mice and humans (Jonckers et al., 2015). In addition, it is difficult to assess exactly how various behavioral tasks commonly used in rodents correlate with clinical assessments, as a majority of the clinical tests rely on language-based communication, which obviously cannot be accurately recapitulated in mice (Webster et al., 2014).

Finally, and perhaps most importantly, genetic diversity has largely been ignored in preclinical studies utilizing mouse models of AD. Currently, most AD models are maintained on only a single or few genetic backgrounds. This includes backgrounds of mixed origins such as the B6SJL background discussed previously, as hybrid strains were traditionally utilized for the generation of transgenic mice due to more desirable reproductive characteristics and enhanced egg quality over traditional inbred strains (Charlesworth and Willis, 2009). This genetic variability may confound interpretation when unaccounted for in studies of genetically diverse individuals. Even when utilizing a single inbred strain, the experimental design is equivalent to studying AD in a single human (or set of identical twins), as each mouse from a given inbred strain is genetically identical. While these study designs are ideal for studying the impact of a specific experimental manipulation and have led to numerous insights and advances in the field of AD, they prohibit increased understanding of how individual genetic variation may influence susceptibility or resilience to disease. Given our increased understanding of how substantially genetics influences AD risk in human populations (Gatz et al., 2006; Lambert et al., 2013), there is a critical need for the field to develop useful AD models that incorporate genetic diversity in a systematic way.

Modeling Genetic Diversity in the Mouse

The importance of including genetic diversity in studies of model organisms has long been appreciated. As such, a large number of resources and approaches have been developed over the previous decades in order to systematically include this important variable in experimental design (Williams and Williams, 2017). These resources range from relatively low genetic complexity, such as panels of single fully inbred strains (Lusis et al., 2016) and reduced complexity crosses derived from closely related substrains (Kumar et al., 2013), to panels that incorporate increasing levels of genetic complexity, including F2 populations derived from 2-4 parents (Damerval et al., 1994; Schadt et al., 2003), advanced intercrosses with more complicated breeding schemes beyond the F2 generation (Darvasi, 1998), and various extensions of these approaches.
(Peirce et al., 2004; Philip et al., 2011). Even more complex resources consisting of strategically outbred mice in which each individual is genetically unique have been recently refined (Churchill et al., 2012), providing opportunities for unprecedented genetic diversity and high mapping precision. The type of experimental population utilized in a given experiment is highly dependent on the behavioral phenotype being evaluated, the degree of genetic complexity desired, as well as the cost and availability of resources (Williams and Williams, 2017).

Consideration of genetic diversity in AD studies

While most AD models are maintained on a single background strain, genetic background has not been completely ignored in preclinical studies. A few previous studies have demonstrated that genetic background has a strong influence on Aβ levels, although these investigations have largely been carried out in studies utilizing panels of single inbred strains or simple F2 crosses between two standard inbred laboratory strains. For example, Lehman and colleagues (2003) transferred an APP transgene and its congenic interval to three inbred mouse strains (B6, DBA/2J, and 129S1/SvlmJ) by repeated backcrossing. All strains initially exhibited similar levels of mutated APP, but strain-specific differences in APP processing and Aβ generation were identified, as evidenced by highly variable levels of C-terminal APP fragments and both brain and blood Aβ40 and Aβ42 (Lehman et al., 2003). Intercrossing of the B6 and 129S1/SvlmJ (129) carrier lines revealed that high levels of Aβ observed in the parental B6 line and resulting F1 population were due to dominantly acting variants on the B6 genetic background. In addition, a pure DBA/2J (D2) background has repeatedly been associated with decreased susceptibility to amyloid deposition (Jackson et al., 2015; Morihara et al., 2014). Similarly, the A/J background has been associated with protection from amyloid deposition and lethality associated with the TgCRND8 transgene (Sebastiani et al., 2006), which drives expression of high levels of mutated APP under the control of the hamster prion promoter (Chishti et al., 2001). Together, these results suggest distinct loci exist within the mouse genome that may alter APP processing in such a way that protects against the pathological accumulation of Aβ plaques.

Studies utilizing F2 populations of mice have attempted to utilize quantitative trait loci (QTL) mapping (Yang et al., 2007) to identify precise regions of the genome modulating susceptibility to Aβ accumulation. In a QTL mapping experiment, quantitative measures of a specific and well-defined trait of interest are measured across a population, and alleles that co-vary with a quantitative trait are identified and denoted as QTL. For example, Ryman and colleagues (2013) generated an F2 population from B6 and D2 mice, each member of which carried an APP-based transgene [R1.40, (Lamb et al., 1993; Ryman et al., 2008)]. Using Aβ40 as their quantitative trait, the group identified several large genomic regions harboring variants that influencing extent of deposition. Sebastiani and colleagues utilized a similar approach, creating an F2 population from B6 and A/J mice harboring the TgCRND8 transgene (Sebastiani et al., 2006). A number of QTLs were identified that modified the total number of deposited plaques as measured by Thioflavin S staining. However, in both of these studies, the
identified QTLs were large and precise genes involved in modifying amyloid deposition were not resolved. In addition, neither study evaluated the impact of genetic background on cognitive deficits associated with AD transgenes. Given the general lack of an association between amyloid levels and cognitive function in human populations (Villemaigne et al., 2011), distinct genetic loci likely regulate susceptibility to amyloid accumulation and cognitive deficits in AD. As such, much remains to be learned in the context of mouse genomic variants and their influence on AD-related cognitive symptoms in transgenic models.

**Recombinant inbred strains**

While F2 crosses are advantageous due to their ability to produce large numbers of mice from almost any existing resource and relative ease of mapping, they also present a number of challenges. For one, F2 mice generally exhibit a relatively low number of recombinations per genome (Williams and Williams, 2017), and as a result QTL mapping typically yields large genomic intervals unless large numbers of animals are utilized. In addition, each F2 mouse is genetically unique, which requires that each individual be genotyped, driving up costs associated with any given experiment. Perhaps most importantly, the utilization of genetically unique, and thus non-reproducible animals, greatly limits the ability for repeated sampling of the same genotype at different time points, in different laboratories, or under different experimental conditions. As such, it is difficult, if not impossible, to build a comprehensive “phenome” of an F2 population (Williams and Williams, 2017) and truly understand how DNA variants may act to influence molecular and behavioral outcomes across a genetically diverse population.

To address a few of these weaknesses associated with standard F2 crosses, recombinant inbred (RI) strains have emerged as one of the most useful resources for evaluating the impact of genetic diversity on behavioral and/or molecular outcomes (Andreuex et al., 2012; Philip et al., 2011; Wang et al., 2016; Williams et al., 2016). In the standard creation of RI strains, genomes from a set of initial parental strains are scrambled through breeding, typically through the F2 generation. Pairs of F2 offspring are then randomly selected to serve as founders for new inbred strains of mice, and sibling inbreeding is repeated at each subsequent generation until a new inbred strain is established at about generation 20 (Silver, 1995). **Figure 1-3** illustrates construction of a set of RI strains from B6 and D2 progenitor strains. Each new inbred strain created (e.g. BXD101, BXD102, and BXD103 in **Figure 1-3**) is referred to as an individual RI strain, as it contains a unique set of recombinations that have become fixed by repeating inbreeding. Practically, this means that each RI strain contains a unique combination of homozygous parental alleles across the genome, providing a resource that systematically incorporates naturally occurring genetic variation. Each of the new inbred strains produced from the same set of parental strains are considered to be members of the same RI “set”, which can then be utilized in the same manner as any other mapping cross to evaluate effect of genetic diversity on a variety of behavioral and molecular traits.
Figure 1-3. Construction of a set of RI strains from B6 and D2 progenitors.

This figure illustrates the process of RI strain development. Mice from two distinct inbred strains (here, B6 is indicated by shaded chromosomes and D2 is indicated by non-shaded chromosomes) are mated together. The F1s produced from that initial crossing are then mated together to produce F2s. Circled genomic regions at the F2 generation indicate genomic segments that have already become fixed by the F2 generation. F2 mice are then sibling inbred for at least 20 generations until new inbred strains (here, BXD101, BXD102, and BXD103) are created. The new strains are referred to as recombinant inbred because they are inbred strains harboring a unique series of recombinations from the initial founder strains.

BXD genetic reference panel

While a number of RI panels have been generated, the set of RI strains resulting from an initial crossing of B6 and D2 progenitors, known as the BXD genetic reference panel (GRP), is currently the largest and best-characterized (Wang et al., 2016). Initially, BXDs were created exactly as described above – inbreeding random pairs of F2 offspring to create new RI lines (Taylor et al., 1999; Taylor, 1978). However, Peirce and colleagues (2004) recognized that in order to maximize the utility of RI strains, a higher density of recombinations than that provided by this standard design would be required (Peirce et al., 2004). To achieve this goal, a second set of BXD RI strains was generated from two separate advanced intercross populations (Darvasi and Soller, 1995) where offspring were strategically bred for up to 10 generations prior to the start of inbreeding (Peirce et al., 2004). As a result, BXD lines exist that contain a higher recombination density than other RI populations, and as such, greater potential mapping precision. Currently, the BXD panel consists of ~150 unique RI lines, each of which have been densely genotyped and are known to segregate for over 5 million single nucleotide polymorphisms (SNPs), around 12,000 of which are coding missense mutations predicted to impact protein structure and/or function (Wang et al., 2016).

Utility for systems genetics

RI strains in general, and the BXD panel in particular, present a variety of advantages over other genetically diverse resources. Perhaps most important is their utility for systems genetic analyses, a branch of systems biology aimed at understanding the flow of biological information that underlies complex traits (Civelek and Lusis, 2014). A foundation of systems genetics is the accurate collection of trait data, which is greatly facilitated by the use of the BXD panel. As each RI line is fully inbred, large numbers of genetically identical individuals can be generated, allowing for the resampling of identical genotypes. This practice reduces non-genetic variance in trait data (Belknap, 1998), providing both additional power and confidence in experimental results and making strides toward addressing the reproducibility crisis that has emerged in recent years (Begley and Ioannidis, 2015). In addition, as each BXD strain has already been genotyped, individuals generated as part of a new experiment do not need to be re-genotyped. As such, QTL maps yielding insight into specific genomic regions harboring variants that regulate complex traits can be quickly and feasibly generated, providing a starting point for large-scale systems genetics analyses.

Systems genetics relies on the fundamental hypothesis that variations in genomic sequence cause variations in behavioral phenotypes through complex molecular cascades involving multiple layers of information, including RNA and protein (Civelek and Lusis, 2014; Mulligan et al., 2017). In order to precisely understand these cascades, systems genetics approaches advocate the sampling of multiple intermediate phenotypes at a global scale, which has been facilitated by advances in various “-omics” technologies, including epigenomics, transcriptomics, proteomics, and even metabolomics. Typically, it is difficult to acquire such a wealth of data from a single sample (or single individual). However, the inbred nature of the BXDs allow for the comparison of multiple scales of
data from the same “individual” without the reliance on a single sample. In addition, as BXDs can be generated across time and laboratories, the burden of generating each scale of data does not fall on a single research group but instead facilitates the collection of thousands of independently collected traits into a comprehensive “phenome”. Currently, the BXD phenotype is stored on GeneNetwork.org (Mulligan et al., 2017) and consists of around 7,300 quantitative metabolic, physiological, pharmacological, and behavioral phenotypes in combination with genome-wide transcriptional, proteomic, or metabolomic data from over 20 unique tissues (Wang et al., 2016). Various approaches ranging from simple correlation analyses to complex statistical modeling can then be used to integrate these various scales of data and generate hypotheses regarding causal mechanisms, ultimately enabling understanding how DNA sequence variation yields complex variation in a number of traits.

Recombinant inbred backcross

Given the significant advantages presented by the utilization of RI strains, it is not surprising that they have been expanded upon in a variety of ways (Williams and Williams, 2017). One of the most interesting uses of RI strains involves generating a set of F1 intercrosses between an RI strain and a single inbred strain, often called an RI backcross (RIB). In this design, the single inbred strain often contains a specific gene variant, mutation, or transgene of interest that confers a specific phenotype in a dominant or additive fashion. The introduction of the genetically diverse chromosome from the RI strain allows the identification of specific modifier genes that alter the expressivity of the mutation or transgene. For example, Hunter and colleagues (2001) first utilized this approach to identify modifiers of cancer metastasis segregating among the AKXD RI strains, an RI panel generated from an initial cross between AKR/J and D2 mice (Hunter et al., 2001). Similarly, Bennett and colleagues (2015) utilized BXD strains, in combination with other RI strains as part of the Hybrid Mouse Diversity Panel, to investigate the genetic architecture of atherosclerosis in mice. They identified a number of genetic risk factors, including some overlapping with previous human studies (Bennett et al., 2015). This experimental design is particularly advantageous, as F1 experimental mice are all isogenic but recombinant, providing a reproducible resource to study effects of transgenic models in the context of genetic diversity without the years of breeding necessary to create a completely new RI panel (Hunter et al., 2001; Peirce et al., 2004). Specifically, the genetic architecture of a RIB resembles a conventional N2 cross, but with a much higher load of recombinations per genome (Hunter et al., 2001), providing opportunities for precise QTL mapping of modifier regions.

Hypothesis and Specific Aims

Given the importance of genetic background in modifying both susceptibility to cognitive decline in aging (Swan et al., 1990), and risk of developing AD (Gatz et al., 1997; Gatz et al., 2006), we hypothesized specific genes involved in modifying cognitive decline in aging and AD could be identified using genetically diverse mouse models.
specifically the BXD GRP. Additionally, given the important role of hippocampal neuronal circuitry in cognitive function (Bliss and Collingridge, 1993; Kaczorowski and Disterhoft, 2009; Kaczorowski et al., 2011), we hypothesized that those genes that most robustly altered susceptibility or resilience to cognitive decline would do so but regulating hippocampal neuronal excitability and/or synaptic plasticity. To directly test this hypothesis, I propose the following specific aims:

**Specific Aim 1: Evaluate the effect of individual genetic differences on age-related cognitive decline.**

While the BXDs have been successfully used to identify genomic regions important for determining learning and memory capabilities early in life (Wehner et al., 1997), how naturally occurring variation across the BXDs affects cognitive aging is not well understood. Chapter 2 describes the cognitive characterization of 21 BXD strains at midlife and the identification of a region on chromosome 4 containing variant(s) that directly modulate cognitive aging.

**Specific Aim 2: Evaluate the effect of individual genetic differences on AD-related cognitive decline and neuropathology.**

In order to better understand how individual genetic variation modifies susceptibility to AD, I developed the first genetically diverse mouse model of AD using a RIB strategy combining the 5XFAD model of AD and the genetically diverse BXD GRP (termed the AD-BXDs, discussed in Chapter 3). We then go on to test the hypothesis that inclusion of genetic diversity would improve the translational validity of our model, and show substantial overlap with human AD at the genetic, transcriptional, and behavioral levels.

**Specific Aim 3: Identify specific genes involved in modifying susceptibility to age and AD-related cognitive decline through QTL mapping and prioritization analyses.**

Identification of modifier loci and predictive gene variants with the potential to act as biomarkers for earlier detection of AD is a critical step toward the development of personalized therapeutics and mechanism-based treatments. Chapter 2 describes the identification, prioritization, and validation of the gene heterochromatin protein 1 binding protein 3 (H3bp3) as a critical regulator of normal age-related cognitive decline. Chapter 4 describes the identification of novel loci modifying susceptibility to AD-related cognitive decline and development of amyloid pathology, as well as the identification of transcriptional networks present early in disease that may predict later-life resilience to symptom development. Together, these results provide both a better understanding of genetic and molecular mechanisms that underlie cognitive decline and a valuable list of candidate genes that may be targeted as therapeutics to prevent or delay the onset of disease.
CHAPTER 2. SYSTEMS GENETICS IDENTIFIES HP1BP3 AS A MODIFIER OF NORMAL COGNITIVE AGING*

Introduction

Aging is associated with a decline in cognitive performance that begins around midlife (i.e. 45–50 years, but the onset and severity varies greatly among individuals (Davies et al., 2015; Singh-Manoux et al., 2012). Although variation in cognitive performance at midlife is highly heritable [i.e. 62–78% of variance is attributable to genetic factors, (McClearn et al., 1997; Plomin and Deary, 2015)], a large portion of this heritability remains unexplained by currently identified gene variants (Johnson et al., 2015). Several studies have identified associations between APOE, BDNF, and catechol-O-methyl transferase with either cognitive ability or rate of cognitive decline in older people (Harris and Deary, 2011; Laukka et al., 2013; Payton, 2009; Tapia-Arancibia et al., 2008; Wisdom et al., 2011). However, even when polymorphisms in these genes are significantly associated with cognitive phenotypes, effect sizes are typically small (Harris and Deary, 2011), indicating that additional genes contribute significantly to the regulation of cognitive decline in human populations. The identification of specific genes that modify the development and progression of cognitive decline may aid in early diagnosis of at-risk patients, as well as identify novel targets for the development of therapeutics to prevent or delay the onset of disease.

Challenges to identifying DNA variants that modify cognitive aging in humans include substantial genetic heterogeneity, difficulty in controlling environmental factors, and limited molecular data from disease-relevant human brain tissue. In order to circumvent some of these challenges, mouse GRPs have been designed to model some of the genetic and phenotypic complexity of human populations (International HapMap et al., 2010; Peirce et al., 2004; Williams et al., 2001), allowing a researcher to exploit phenotypic heterogeneity across a population while controlling for environmental factors (Williams and Auwerx, 2015). One such model, a well-characterized GRP known as the BXDs (Peirce et al., 2004; Taylor, 1978; Taylor et al., 1999), was derived by crossing two common inbred strains, B6 and D2. The BXDS have been successfully used to identify genomic regions important for determining learning and memory capabilities early in life (Wehner et al., 1997), but have not yet been used to study cognitive aging.

Here we perform a forward systems genetic analysis on an aged cohort of strains from the BXD GRP.

**Methods**

**Animals**

Male and female mice were group housed (2–5 per cage) and maintained in colony housing (12-hour light/dark cycle) with ad libitum access to food and water. Middle-aged mice (15 ± 0.3 months, 2–8 mice per strain) from 21 BXD RI strains were phenotyped for mapping studies. *Hnis1bp3^2m1a(EUCOMM)Wtsi* mice (n = 6/group) were acquired from the European Conditional Mouse Mutagenesis program and are described elsewhere (Garfinkel et al., 2015a). For viral experiments, male 3-6 month-old B6 and D2 mice were obtained from The Jackson Laboratory (JAX #000664 and #000671, respectively). Mice were group housed and maintained on a 12 hour light/dark cycle with ad libitum access to food and water. Mouse experiments were conducted according to the NIH Guide for Care and Use of Laboratory Animals. All procedures and protocols were reviewed and approved by The Animal Care and Use Committee of The Jackson Laboratory, the UTHSC Animal Care and Use Committee, and/or the Institutional Committee on Animal Care and Use at the Hebrew University of Jerusalem where appropriate.

**Behavioral assays**

**Contextual fear conditioning**

All mice were habituated to transport for at least three days prior to behavioral tests. On the fourth day, mice were trained on a standard CFC paradigm (Neuner et al., 2015). Briefly, CFC consisted of a 180 s baseline period followed by 4 mild foot shocks (1 s, 0.9 mA, inter-shock interval 115 ± 20 s). The average activity burst following each shock was determined for each by manual timing and used to quantify the post-shock reactivity for each mouse to test for possible strain-specific differences in pain sensitivity. Twenty-four hours later, hippocampus-dependent CFM (Chen et al., 1996) was tested in a 10 min session. Behavioral freezing, an index of conditioned fear, was measured using Freeze Frame software (Coulbourne Instruments, PA) at the University of Tennessee or EthoVision software (Noldus Information Technology, Netherlands) at the Hebrew University of Jerusalem. Our lab has previously demonstrated that measures of freezing obtained via video monitoring software correlates well with hand-scored measures of freezing (Kaczorowski et al., 2012). Experimenters were blind to the genotype during behavioral tests and data analysis.
T-maze

The T-maze test was performed as described previously (Deacon and Rawlins, 2006). Mice were given 10 minutes to habituate to the testing room. The T-maze floor was coated with fresh woodchips before each round of trials. Mice were placed into the T-maze and allowed to choose one of two “goal” arms. The mouse was then confined to that arm for 30 seconds using a central partition. The mice were removed from the maze and immediately placed back at the starting point, the partition removed, and the mouse allowed to choose between the two goal arms again (Deacon and Rawlins, 2006). An “alternation” occurred when the mouse entered the opposite arm than that to which it was just confined. The criterion point used was whole animal in goal arm, including tail tip. A total of six trials were performed, with animals tested no more than twice a day. The alternation scores across all six trials were averaged to obtain an average alternation score for each mouse. For viral experiments, working memory was assessed 4-6 weeks following viral injections and were habituated to the testing room for three days prior to testing.

Calculation of memory index for adult BXD strains

BXD strains aged 8–9 weeks had previously been subjected to CFC by Philip and colleagues (2010). The dataset used in our study (including males and females) is publicly available on GeneNetwork.org as BXD Published Phenotypes Record ID 11908. The memory index used in Trait 11908 is a measure of activity (beam breaks) throughout CFM testing. This measure of activity has been previously shown to correlate well with measures of hand-scored freezing behavior (Valentinuzzi et al., 1998). As the dynamic range of human-observed freezing behavior is 0–100% (Anagnostaras et al., 2000), and an activity score of zero is equivalent to 100% freezing, we subtracted these activity measures from 100 to obtain a measure of inactivity that parallels the CFM index (% freezing) calculated for middle-aged strains. Specifically, the mean time each adult BXD strain spent active during CFM testing was subtracted from 100 to obtain a measure of inactivity. The genotype of each strain at the Hp1bp3 locus was then obtained from GeneNetwork.org and used for genotype-by-memory analyses.

QTL identification and candidate gene selection

After CFM tests for 21 BXD strains (mean age 15 ± 0.3 months, 2–8 mice per strain) were completed, the mean time spent freezing was averaged for each strain. Strain means were entered in GeneNetwork.org and are publicly available as BXD Published Phenotypes Record ID: 18395. Hippocampal transcript data from aged BXD strains available on GeneNetwork.org as UTHSC BXD Aged Hippocampus Affy MoGene1.0ST (May15) RMA Gene Level, were used to identify cis (locally)-regulated genes (Andreux et al., 2012; Chesler et al., 2005) from within intervals of interest. The same transcript data was used to identify genes whose expression correlated with memory function in aged animals. Only probes that did not overlap single nucleotide polymorphisms were
used in order to avoid hybridization artifacts that may cause apparent differences in expression due to technical rather than biological variance. When multiple SNP-free gene level probes were available (e.g. *Hp1bp3*), the mean expression of exon-targeting probes was used. In cases where SNP-free gene level probes were not available, exon-level hippocampal transcript data from aged BXD strains was used (UTHSC BXD Aged Hippocampus Affy Mouse Gene1.0ST (Sep12) RMA Exon Level, dataset accession ID GN392). If exon level probes were used, the probe with highest average expression across all strains was selected. All probes were verified by BLAT (UCSC Genome Browser). Correlation analyses for initial candidate gene prioritization used Spearman’s rho and were adjusted for multiple comparisons using Bonferroni correction (95% confidence interval).

Western blots

**Mouse samples**

Whole hippocampal lysates were prepared as previously described (Neuner et al., 2016). Protein concentration was determined using a NanoDrop Spectrophotometer 2000 and 20 µg total protein was loaded onto a BioRad TGX Mini Protean gel. Proteins were separated by electrophoresis and transferred to a nitrocellulose membrane using the Trans-Blot Turbo system (BioRad). Membranes were blocked in 5% non-fat milk in PBST and primary antibodies for HP1BP3 (gifted by Drs. Benjamin Garfinkel and Joseph Orly) and GAPDH (Fitzgerald Industries #10R-G109A) were incubated at 4°C overnight. Fluorescently labeled secondary antibodies were incubated for 1 hour at RT and bands were visualized using an Azure Biosystems gel imager. Observed double band staining is typical expression pattern for HP1BP3 (Garfinkel et al., 2015b) and overlaps with positive control HP1BP3 overexpression lysate from human 293T cells (Abnova #H0005809-T02), which was used as a positive control. As such, both bands were included in our analysis. Band densities were measured using LI-COR Image Studio Lite (LI-COR Biosciences). Densities were first normalized to an internal B6 control, and then total HP1BP3 was adjusted for total protein loaded using GAPDH. Adjusted densities are plotted in appropriate figures.

**Human samples**

Frozen human hippocampal samples (n=3/group) collected postmortem were obtained from the University of Kentucky Sanders-Brown Center on Aging and stored at −80°C until use. Briefly, hippocampal lysates were prepared from frozen tissue, protein concentration was determined using a Nanodrop2000 Spectrophotometer (ThermoScientific), and 20 µg of total protein was loaded and separated on a 10% SDS-PAGE gel. Proteins were transferred using the Bio-Rad TurboTransfer system and blocked for 30 min at room temperature. Primary antibodies for HP1BP3 and GAPDH (ProteinTech #24556-1-AP and Fitzgerald #10R-G109a, respectively) were incubated overnight and detected by anti-mouse and anti-rabbit fluorescent conjugated antibodies. Visualization was performed using an Odyssey image scanner and blots were quantified.
using the Odyssey software version 5.0 (LiCOR). Results were replicated in 2 independent Western blots. Observed double band staining is typical expression pattern for HP1BP3 (Garfinkel et al., 2015b) and overlaps with positive control HP1BP3 overexpression lysate from human 293T cells (Abnova #H00050809-T02), which was used as a positive control. Loading-dye only lanes served as negative control.

**Transcriptome analyses**

**RNA sequencing**

Following viral experiments, total RNA was prepared from whole hippocampal homogenates using the Qiagen miRNeasy Mini kit (#217084). RNA quality was assessed using a BioAnalyzer and only samples with RNA Integrity Number (RIN) greater than 8.0 were included. Samples were submitted to The Jackson Laboratory Genome Technologies department, where mRNA libraries were prepared using the KAPA Biosystems mRNA Hyper Prep Kit while miRNA libraries were prepared using the Illumina TruSeq Small RNA Library Prep Kit. The Illumina HiSeq 2500 was used to sequence 75 paired-end reads, at a depth of 30 million reads for mRNA and 10 million reads for miRNA. Two pairs of mRNA/miRNA samples given identical barcodes during the initial sequencing, so all samples were re-sequenced and where available, fastq files from both runs were concatenated, providing expanded sequencing depth for 10/12 samples. Poor quality mRNA reads were filtered, trimmed, and aligned to either the B6 or D2 genomes where appropriate using rsem v1.2.12, followed by bowtie2 v2.2.0 for expression estimate of trimmed reads (Li and Dewey, 2011). For miRNA, trimmed reads were aligned to the mouse miRNA-precursor miRBase reference (miRBase release 21) using bowtie v1.0.0 (Griffiths-Jones, 2006; Kozomara and Griffiths-Jones, 2014). Following alignment miRNA counts were extracted using featureCounts function from the subread package v1.5.2 (Liao et al., 2013).

**Differential expression analysis**

For both RNA types, expected read counts were used to perform differential expression analysis using DESeq2 according to established protocols (Love et al., 2014). miRNAs of interest were checked for single nucleotide polymorphisms (SNPs) using the Sanger Mouse Genomes database (Keane et al., 2011) to ensure observed differences were not due to strain-specific differences in miRNA sequence. For graphing of mRNA and miRNA differences across groups, normalized counts were extracted from DESeq2 using the counts function. miRNA counts were further log transformed using the rlog function.

**Gene set enrichment analyses**

For analyses utilizing data obtained from GeneNetwork, the following protocol was utilized: Briefly, all nominally significant correlates of hippocampal Hp1bp3 were extracted from whole-genome hippocampal transcript data from BXD strains age-
matched to those used for CFC. Correlates were ranked based on correlation coefficient and this preranked list was used to calculate an enrichment score for gene sets obtained from the Molecular Signatures Database (MSigDB) version 5.0. Enrichment scores are obtained by calculating a cumulative ("running-sum") statistic, which is increased when a gene is present in the set being tested and decreased when a gene is not. The maximum deviation of this cumulative score from zero is the enrichment score. Then, normalized enrichment scores (NES) were obtained to account for the size of the gene set being tested. The proportion of false positives was controlled using established methods (Mootha et al., 2003; Subramanian et al., 2005). Based on these criteria, any gene set with a false discovery rate (FDR) of ≤ 0.25 was accepted as significantly enriched in our dataset.

For analyses using RNA sequencing data generated from mice undergoing viral experiments, the following protocol was utilized: All genes nominally significantly differentially expressed relative to treatment (Hp1bp3 KD vs Ctrl unadjusted p-value < 0.05) were sorted into a ranked list by log2 fold change, with the most positively differentially expressed genes at the top of the list. This pre-ranked gene list was uploaded into the Gene Set Enrichment Analysis (GSEA) desktop software (Subramanian et al., 2005). GSEA was used according to established procedures to identify significant overlap between genes positively or negatively changed by treatment and gene sets in MSigDb 3.0 (Liberzon et al., 2011). The same list of differentially expressed genes was uploaded into Ingenuity Pathway Analysis [IPA, Qiagen, (Kramer et al., 2014)] and the Core Analysis function was utilized to identify upstream miRNA regulators of observed mRNA changes. Specifically, we filtered results to include references from the Ingenuity Knowledge Base, included both direct and indirect connections, and considered only molecules and/or relationships that had been experimentally observed or had high predictive value. We then filtered the putative upstream regulators list to only include regulators of molecule type equal to “miRNA” or “mature miRNA”. Finally, to perform simple over-representation enrichment analysis on small gene sets (e.g. miRNA target genes) as opposed to searching specifically for enrichment among up or down-regulated genes (as in GSEA), we utilized WebGestalt (Wang et al., 2013) to identify significantly enriched GO Biological Process terms.

Design of viral vectors

Adeno-associated viral (AAV) vectors were made by Vector Biolabs (Philadelphia, PA, USA). shRNA targeting Hp1bp3 was packaged into AAV serotype 9 (AAV9) vector and stored in a 1X PBS buffer containing 0.001% pluronic F-68, 0.22 mm filter sterilized. A scrambled shRNA control was also synthesized according to identical specifications. The resulting vectors were subjected to purification by 2 rounds of CsCl Density Gradient Centrifugation combined with ultracentrifugation to separate empty and full capsids.

26
**Intrahippocampal injections**

Delivery of viral vectors into the hippocampus was performed as previously described (Neuner et al., 2015). Briefly, mice were anesthetized under 4-5% inhalable isoflurane gas and once anesthetized, positioned in a stereotaxic frame and maintained at 1-2% inhalable isoflurane gas. Guide cannula and a 28-gauge microinjector attached to a Hamilton syringe and syringe pump were used to deliver 1.0 uL of virus per hemisphere at a rate of 1.0 uL/min. Microinjectors were left in place for 5 minutes following the end of injection to allow virus to diffuse adequately. Following injection, bone wax was applied to the skull and the cut was sutured using absorbable sutures. Post-surgery, topical bupivacaine was administered and mice were monitored once daily for a minimum of three days.

**Statistical analysis and data availability**

All experiments and analysis were conducted with experimenters blind to treatment group. All statistics were performed using SPSS (IBM) or R and tests used included independent t-tests, one- and two-way ANOVA, two-way repeated measures ANOVA, and Pearson correlation with confidence levels set a p < 0.05. All analyses were corrected for multiple comparisons and described in the Results section. Unless otherwise stated, data values reported here are given as mean ± SEM. All raw and processed data has been deposited to Gene Expression Omnibus (GEO) and is available as SuperSeries GSE119321. Specifically, the mRNA sequencing data is available as GSE119318 and the miRNA sequencing data is available as GSE119319.

**Results**

**Genetic mapping identifies an interval on chromosome 4 associated with memory status at midlife**

To identify genes involved in the regulation of cognitive aging, we analyzed hippocampus-dependent memory function across a cohort of middle-aged mice (15 ± 0.3 m) that model the genetic and phenotypic variation of human populations (the BXD GRP, **Figure 2-1A**). Specifically, mice were trained on standard CFC (Kaczorowski and Disterhoft, 2009). During training on day 1, mice received 4 scrambled foot shocks in the conditioning chamber over a 10-minute training session. Twenty-four hours later, mice were returned to the original conditioning chamber and the percentage of time each mouse spent freezing over the 10-minute test was measured as an index of CFM. CFM is highly variable across this aged BXD family (21 strains tested, n = 2–8 mice/strain, **Figure 2-1B**) and heritability estimates that compare the observed genetic (between-strain) variance to technical (within-strain) variance (heritability, $h^2 \approx 0.7$) demonstrate that much of this variability is attributable to genetic factors. Although strain-specific
Figure 2-1. Identification of \( Hp1bp3 \) as top candidate modulator of cognitive aging.

(A) The BXD panel was derived by inbreeding the F2 progeny of a B6 and D2 intercross. (B) CFM varies widely across 21 BXD strains (n = 2–8/strain, mean age = 15 ± 0.3 months). Memory index was quantified by percent time spent freezing during 10 min test. (C) Genetic interval mapping revealed a significant CFM QTL on mouse chromosome 4 (Chr4: 137.5–140.5 Mb). Pink horizontal line: Genome-wide statistical significance (\( p = 0.05 \)), Green additive effect line: \( D \) parental allele increases trait values, red: \( B \) allele increases trait values. Yellow tick marks: presence of genome sequence variant. (D) \( Hp1bp3 \) expression is highly correlated with CFM across BXD strains (Pearson \( r = 0.6, p < 0.05 \)), with strains inheriting the \( D \) allele having lower levels of hippocampal \( Hp1bp3 \) and poorer CFM.
differences in acquisition were also observed, the variance in acquisition was not sufficient to explain the differences in CFM across the panel.

Subsequent interval mapping using the average memory index for each strain highlighted a region of chromosome 4 (Chr 4, 137.7–140.5 Mb) that was significantly associated with variation in CFM in middle-aged mice (Figure 2-1C). This QTL had not associated with memory function in studies using younger adult sex-matched BXDs (Philip, et al., 2010), suggesting the QTL contains variants that contribute to variation in cognitive aging as opposed to general memory function.

Prioritization of positional candidates identifies Hp1bp3 as putative regulator of cognitive aging across BXD panel

In order to identify genes located in the QTL that may be causally involved in regulating CFM abilities at midlife, positional gene candidates were prioritized based on: 1) annotated sequence differences segregating among the BXD strains, 2) local control of gene expression as determined by expression QTL analysis using hippocampal transcriptome data from aged BXD strains, and 3) significant correlation between hippocampal gene expression and CFM. The gene Hp1bp3 emerged as the single best positional candidate (Table 2-1).

Hp1bp3 contains multiple missense variants in coding regions, numerous non-coding variants, and insertions/deletions predicted to impact protein function, transcriptional regulation and/or splicing (McLaren et al., 2010; Supplemental Table 1). Hippocampal Hp1bp3 expression was significantly correlated with CFM status, with those BXD strains inheriting the D parental allele exhibiting lower levels of Hp1bp3 transcript and worse CFM performance relative to the B allele (Figure 2-1D). While Hp1bp3 genotype had no effect on memory status in younger adult mice (age of 8–9 weeks, Figure 2-2A, secondary analysis of data from Philip and colleagues, 2010), strains inheriting the D allele exhibited significantly impaired CFM at midlife (Figure 2-2B). Since variation in the CFM index at midlife was not due to genetic differences in baseline anxiety or post-shock pain sensitivity confounds (Figure 2-3), these data suggest a genotype-by-age interaction in which reductions in Hp1bp3 expression correspond specifically to aging-related cognitive impairment (Figures 2-1 through 2-3). In support, the genotype at Hp1bp3 can account for as much as 52% of the strain variance in CFM performance at midlife. However, calculations of genetic variance explained by a QTL are biased upward by an average of 30% due to a combination of epistatic, genotype-by-environment interactions, and small sample size (n < 1000, Beavis, 1998; Wellenreuther and Hansson, 2016; Wurschum and Kraft, 2014). Therefore, we estimate that the actual biological heritable variation in CFM at midlife explained by Hp1bp3 genotype is more likely in the range of 20-25%.
<table>
<thead>
<tr>
<th>Gene Symbol</th>
<th>Description</th>
<th>SNP Count</th>
<th>Trait ID in Gene Network</th>
<th>Location</th>
<th>Max LRS</th>
<th>CFM Corr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hp1bp3</td>
<td>Heterochromatin protein 1 binding protein 3</td>
<td>106</td>
<td>10338180*, 10344447*, 10342675*</td>
<td>Chr 4: 137.7</td>
<td>Chr4: 139.3</td>
<td>r = 0.7, p &lt; 0.05</td>
</tr>
<tr>
<td>Eif4g3</td>
<td>Eukaryotic translation initiation factor 4 gamma 3</td>
<td>763</td>
<td>10509463*</td>
<td>Chr4: 137.5</td>
<td>Chr4: 140.2</td>
<td>r = 0.5, N.S.</td>
</tr>
<tr>
<td>Capzb</td>
<td>Capping protein (actin filament) muscle Z-line, beta</td>
<td>448</td>
<td>10509620, 10509628*</td>
<td>Chr4: 138.7</td>
<td>Chr4: 138.2</td>
<td>r = 0.4, N.S.</td>
</tr>
<tr>
<td>Ifo2</td>
<td>Intermediate filament family orphan 2</td>
<td>225</td>
<td>10509777, 10509781*</td>
<td>Chr4: 139.1</td>
<td>Chr4: 139.3</td>
<td>r = -0.4, N.S.</td>
</tr>
<tr>
<td>Padi2</td>
<td>Peptidyl arginine deiminase, type II</td>
<td>132</td>
<td>10509838, 10509846*</td>
<td>Chr4: 140.5</td>
<td>Chr4: 140.5</td>
<td>r = 0.4, N.S.</td>
</tr>
<tr>
<td>Kif17</td>
<td>Kinesin family member 17</td>
<td>119</td>
<td>10509526, 10509532*</td>
<td>Chr4: 137.8</td>
<td>Chr4: 138.2</td>
<td>r = -0.1, N.S.</td>
</tr>
<tr>
<td>Pla2g2f</td>
<td>Phospholipase A2, group IIF</td>
<td>40</td>
<td>10517646, 10517654*</td>
<td>Chr4: 138.3</td>
<td>Chr4: 140.2</td>
<td>r = 0.2, N.S.</td>
</tr>
<tr>
<td>Mrto4</td>
<td>MRT4, mRNA turnover 4 homolog (S. Cerevisiae)</td>
<td>32</td>
<td>10517706, 10517711*</td>
<td>Chr4: 138.9</td>
<td>Chr4: 139.9</td>
<td>r = -0.5, N.S.</td>
</tr>
<tr>
<td>Nb1l</td>
<td>Neuroblastoma, suppression of tumorigenicity 1</td>
<td>10</td>
<td>10517677, 10517681*</td>
<td>Chr4: 138.6</td>
<td>Chr4: 138.2</td>
<td>r = -0.5, N.S.</td>
</tr>
<tr>
<td>Camk2n1</td>
<td>Calcium/calmodulin-dependent protein kinase II inhibitor 1</td>
<td>4</td>
<td>10509568, no SNP-free probe</td>
<td>Chr4: 138.0</td>
<td>Chr4: 140.5</td>
<td>--</td>
</tr>
</tbody>
</table>

Genes located under the significant CFM QTL were first prioritized for presence of annotated sequence variants between parental strains B6 and D2. Whole-genome transcript data from aged BXD strains was then used to perform expression QTL mapping and those genes whose expression had a peak log ratio statistic (LRS) at the location of the gene (cis, or locally, regulated genes) were given higher priority. These genes were tested for correlation with the CFM trait using SNP-free probes from gene-level analyses (or exon-level analyses where necessary, see Methods section). Spearman’s rho correlations are reported. *Denotes SNP-free probes; N.S. = not significant.
Figure 2-2. Effect of *Hp1bp3* genotype on CFM status is age-dependent. (A) Memory index for adult BXDs (age = 8–9 weeks) was calculated using CFM data generated by Philip and colleagues (2010) and is publicly available on GeneNetwork.org (Trait ID 110908, see Methods section). Parental origin of *Hp1bp3* allele did not significantly affect memory status in adulthood (*D* allele memory index = 81.0 ± 1.3, *B* = 84.3 ± 1.3, *t*(54) = −1.86, *p* = 0.068). (B) Middle-aged (age = 15.0 ± 0.3 months) BXD strains inheriting the *D* parental allele performed significantly worse on CFM tests than strains inheriting the *B* allele (*D* = 4.8 ± 1.0, *B* = 41.5 ± 20.7, *t*(14) = −4.6, *p* < 0.001). Two-way ANOVA revealed a significant interaction between genotype and age [F(1,66) = 33.5, *p* < 0.001], suggesting that *Hp1bp3* genotype influences memory status in an age-dependent manner.
Figure 2-3. Differences in CFM were not attributable to baseline freezing or post-shock reactivity.

There were no significant effects of strain on either A) baseline freezing [$F(20,50) = 1.009, p = 0.47$] or B) length of post-shock activity burst [$F(20,49) = 1.668, p = 0.07$], indicating no strain-specific differences in baseline fear/anxiety or pain sensitivity, respectively, across 21 BXD strains tested.
Genetic correlation analyses suggest important functional roles for \textit{Hp1bp3} in cognitive aging

There is evidence that groups of highly correlated genes are likely to play a similar biological function and/or act as part of the same biological pathway (Eisen et al., 1998; Zhang and Horvath, 2005). To gain additional insight into mechanisms through which \textit{Hp1bp3} may act to regulate cognitive abilities at midlife, genes whose hippocampal expression significantly correlated with that of \textit{Hp1bp3} were identified using whole-genome hippocampal transcript data from middle-aged BXD strains. All nominally significant correlates (uncorrected \(p \leq 0.05\), 2074 genes) were sorted into a ranked list, with the most highly positively correlated genes at the top (Supplemental Table 2). This list was then used for GSEA (Broad Institute; Subramanian et al., 2005) and tested against gene sets publicly available from Broad Institute’s MSigDB, including Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (Kanehisa and Goto, 2000), Reactome pathways (Matthews et al., 2009), and Gene Ontology (GO) classification (Harris et al., 2004), using a false discovery rate of \(\leq 0.25\) to identify significantly enriched gene sets per standard methods (Subramanian et al., 2005).

Functions of \textit{Hp1bp3} that have been identified include regulation of chromatin structure (Dutta et al., 2014), gene expression (Garfinkel et al., 2015b), cell cycle progression (Dutta et al., 2014), and insulin signaling (Garfinkel et al., 2015a), several of which were supported by our analyses. For example, positive correlates of \textit{Hp1bp3} were significantly enriched both in nuclear localization (Figure 2-4A) and insulin signaling (Figure 2-4B). Additionally, we found \textit{Hp1bp3} gene correlates enriched for terms related to plasma membrane localization and for functions related to neuronal function and excitability, including voltage-gated channel activity, ion channel activity, and G-protein coupled receptor activity (Figure 2-4C). Together, these results support the conclusions suggested by our RNA sequencing data; that \textit{Hp1bp3} may critically influence expression of ion channels and receptors located in the plasma membrane that mediate neurogenesis, neuronal excitability, and plasticity, mechanisms that are disrupted in hippocampal neurons of aged rodent models with cognitive deficits (Burke and Barnes, 2006; Kaczorowski and Disterhoff, 2009; Mattson and Magnus, 2006). Finally, \textit{Hp1bp3} gene correlates also show significant overlap with genes differentially expressed between Alzheimer’s disease patients and controls (Blalock et al., 2004); Figure 2-4D), suggesting the role of \textit{Hp1bp3} in cognitive aging may extend into the regulation of cognitive deficits observed in AD.

Functional validation in a knock-out mouse model confirms novel role for \textit{Hp1bp3} in cognitive function

To directly assess the functional consequence of loss of \textit{Hp1bp3} on hippocampus-dependent learning and memory, we employed a reverse genetics approach. During CFM training, \textit{Hp1bp3} knock-out (KO) and wild-type (WT) mice (Garfinkel et al., 2015a)
Figure 2-4. Gene Set Enrichment Analysis (GSEA) suggest important functional roles for *Hp1bp3* in cognitive aging.
All nominally significant hippocampal correlates (uncorrected p ≤ 0.05, n = 2074, Supplemental Table 2) of *Hp1bp3* were extracted from whole-genome transcriptome data from BXD strains age-matched to those used for CFC. The genes were sorted based on correlation coefficient, with the most highly positively correlated genes at the top of the list. Graphs illustrate calculation of enrichment score, or deviation from random distribution throughout the preranked list. Each vertical black line below the graph represents a gene set member and its position in the preranked list (red=genes positively correlated with *Hp1bp3*, blue=genes negatively correlated with *Hp1bp3*). A FDR cut-off of *q* ≤ 0.25 was used to identify significantly enriched sets according to established methods. (A, top) Positive correlates were significantly enriched for localization to the nucleus. (A, bottom) Genes localized to the nucleus were also present in significantly enriched functional gene sets including RNA splicing, nuclear transport, chromosome organization, cell cycle, and transcription cofactor activity. Normalized enrichment score (NES) represents enrichment score after adjustment to account for size of the given gene set. (B) Positive correlates were significantly enriched for the KEGG pathway corresponding to insulin signaling. (C, top) Negative correlates of *Hp1bp3* were significantly enriched for localization to the plasma membrane. (C, bottom) Genes localized to the plasma membrane were also present in significantly enriched functional gene sets including ion channel activity, substrate-specific channel activity, transmembrane transporter activity, voltage-gated channel activity, and G-protein coupled receptor activity. (D) Correlates of *Hp1bp3* also show significant overlap with a set of genes determined by Blalock and colleagues (2004) to be significantly differentially expressed between AD patients and corresponding controls.
A. GO Cellular Compartment: Nucleus

Gene Set | NES
---|---
RNA splicing | 2.05*
Nuclear Transport | 2.04*
Chromosome organization | 2.00*
Cell Cycle | 1.92*
Transcription co-activator activity | 1.84*

B. KEGG Insulin Signaling Pathway

Gene Set | NES
---|---
(pos. correlated with Hpx1p/3) | (neg. correlated with Hpx1p/3)

C. GO Cellular Compartment: Plasma Membrane

Gene Set | NES
---|---
Ion channel activity | -2.11*
Substrate-specific channel activity | -2.09*
Transmembrane transporter activity | -1.89*
Voltage-gated channel activity | -1.72*
G-protein coupled receptor activity | -1.71*

D. Genes down-regulated in Alzheimer’s disease

Gene Set | NES
---|---
(pos. correlated with Hpx1p/3) | (neg. correlated with Hpx1p/3)

NES, Normalized Enrichment Score, *q<0.25
showed no differences in baseline activity (Figure 2-5A) or shock reactivity (Figure 2-5B), and KO mice exhibited comparable acquisition of conditioned fear, indicating they successfully learned the context-shock association (Figure 2-5A). Nevertheless, KO mice exhibited long-term CMF deficits when tested 24 hours later (Figure 2-5C). Notably, the effect of Hplbp3 KO on CMF was similar to the effect of genotype at the Hplbp3 locus. As the KO mice were generated on a B6 background, the WT mice carry the B version of the Hplbp3 allele, and as expected, perform comparable to B6 mice from the BXD GRP. However, KO mice performed more similarly to BXDs harboring the D allele, which exhibited a 20–40% reduction in conditioned freezing during CMF tests compared to strains with the B allele (Figure 2-2B), and suggests the D allele functions similarly to a loss-of-function mutation.

To determine if loss of Hplbp3 mimics aging-related effects on additional hippocampus-dependent cognitive domains (Zornetzer et al., 1982), working memory was assessed using the T-maze test of spontaneous alternation (Deacon and Rawlins, 2006). KO mice were significantly impaired, and like aged mice (Lalonde, 2002), performed at chance levels (Figure 2-5D). Taken together, these results demonstrate, for the first time, that Hplbp3 is necessary for successful hippocampus-dependent memory function spanning multiple cognitive domains - and suggest that therapeutic interventions to restore levels of Hplbp3 may improve cognitive function.

**Hippocampus-specific knockdown of Hplbp3 impairs cognitive function**

As Hplbp3 KO littermates appear smaller than their WT littermates and have alterations in insulin signaling (Garfinkel et al., 2015a), additional experiments were necessary to establish that alterations in cognitive function were due to central rather than peripheral mechanisms. Specifically, we designed an AAV9 viral vector encoding either shRNA for Hplbp3 or a scrambled control and delivered 1.0 ul. bilaterally to the dorsal hippocampus of adult (3-6 months) B6 and D2 mice (Figure 2-6A). Efficacy of the viral vector to reduce HP1BP3 levels was confirmed via western blot (Figure 2-6B). Notably, no reduction in HP1BP3 levels was observed in the cortex or cerebellum, indicating minimal viral spread outside our target region (Figure A-1). Four to six weeks following injection, we assessed both working and long-term CMF in these mice. There was a significant effect of Hplbp3 knockdown (KD) on working memory as measured in the T-maze, with both B6 and D2 mice receiving AAV9-shRNA-Hplbp3 performing worse than their control counterparts [Figure 2-6B, two-way ANOVA, effect of treatment F(1, 41) = 5.9, p = 0.02]. Long-term CMF was then assessed via CFC. Although D2 mice overall performed more poorly than B6 mice during training, there was no effect of Hplbp3 KD on contextual fear acquisition (CFA) as measured by freezing during the 40s following the final shock [Figure 2-6C, left, two-way ANOVA, effect of strain F(1, 39) = 5.4, p = 0.03, effect of treatment F(1, 39) = 0.12, p = 0.73]. In contrast, there was a significant main effect of treatment on CMF [Figure 2-6C, right, two-way ANOVA, effect of strain F(1, 39) = 56.4, p < 0.001, effect of treatment, F(1, 39) = 13.6, p = 0.001], suggesting Hplbp3 is uniquely involved in mechanisms underlying memory.
Figure 2-5.  *Hp1bp3* KO mice exhibit impairment on hippocampus-dependent long-term and working memory tasks.

(A) Comparable baseline (BL) freezing and CFA by the final trial in WT and *Hp1bp3* KO mice (*n* = 6/group) during CFC. (B) There were no differences in post-shock reactivity during CFC training between WT and *Hp1bp3* KO mice, indicating no differences in pain sensitivity. Videos were manually analyzed to determine the average length of activity burst for each mouse. WT = 2.6 ± 0.2 s, KO = 2.6 ± 0.2 s; *t*(10) = −0.11, *p* = 0.91. (C) KO mice were significantly impaired on CFM [*t*(10) = 3.10, *p* < 0.001]. (D) KO mice also exhibited spatial working memory deficits [*t*(12) = 5.095, *p* < 0.001].
Figure 2-6. Hippocampus-specific knockdown of Hp1bp3 impairs cognitive function.

(A) 1.0 uL of AAV9 encoding either Hp1bp3 shRNA or a scrambled control was injected bilaterally into the dorsal hippocampus of adult B6 and D2 mice. After 4-6 weeks, a subset of mice was behaviorally tested. A minimum of two weeks later, mice were either harvested, and tissue biobanked, or used for electrophysiology. Mouse brain image adapted from The Mouse Brain in Stereotaxic Coordinates (Paxinos and Franklin, 2004).

(B) Left, efficacy of the viral vector to reduce HP1BP3 levels in the hippocampus was confirmed via Western blot; n = 6/group, two-way ANOVA effect of strain F(1, 20) = 1.2 p = 0.3, effect of treatment F(1, 20) = 15.6, p = 0.001, no interaction. Right, data from four independent blots (two experiments) was pooled across strains (n = 12/group) and quantified. Lysate from human 293T cells overexpressing HP1BP3 was used as a positive control. (C) Working memory was measured in the T-maze. A significant effect of treatment was observed F(1, 41) = 5.9, p = 0.02, indicating that Hp1bp3 KD impairs working memory function across B6 and D2 mice. (D) CFA and CFM were assessed using CFC. Although B6 and D2 mice acquired the task differently on Day 1 [effect of strain, F(1, 39) = 5.4, p = 0.03], Hp1bp3 KD had no effect on the extent of these differences [effect of treatment, F(1, 39) = 0.12, p = 0.73]. Right, when mice were placed back in the training chamber on Day 2, D2 mice performed significantly worse than B6 mice [effect of strain F(1, 39) = 56.4, p < 0.001] – so poorly that further impairment caused by Hp1bp3 reduction could not be observed. However, Hp1bp3 KD did impair CFM in B6 mice [effect of treatment, F(1, 39) = 13.6, p = 0.001, interaction between strain and treatment F(1, 39) = 14.5, p < 0.001]. *main effect of treatment, p < 0.05, **t-test, p < 0.05
A. 

**Hp1bp3 shRNA construct**

[Diagram showing the structure of the shRNA construct with ITR, U6, shRNA, CMV promoter, and eGFP.

**Scrambled shRNA control construct**

[Diagram showing the structure of the scrambled shRNA control construct with ITR, U6, shRNA, CMV promoter, and eGFP.]

<table>
<thead>
<tr>
<th>Injections</th>
<th>Behavior</th>
<th>Electrophysiology or harvest</th>
</tr>
</thead>
<tbody>
<tr>
<td>~ 3-6 months</td>
<td>+ 4-6 weeks</td>
<td>+ 2 week min.</td>
</tr>
</tbody>
</table>

C57BL/6J & DBA/2J male mice

B. 

**HP1BP3 Protein**

[Diagram showing Western blot analysis of HP1BP3 and GAPDH with lanes for B6 Ctrl, B6 KD, D2 Ctrl, and D2 KD.]

C. 

**Working Memory**

[Box plot showing percentage alternation for B6 Ctrl, B6 KD, D2 Ctrl, and D2 KD.

D. 

**Day 1: Training**

[Graph showing percentage freezing for B6 Ctrl, B6 KD, D2 Ctrl, and D2 KD.

**Day 2: Testing**

[Graph showing percentage freezing for B6 Ctrl, B6 KD, D2 Ctrl, and D2 KD.]

---

39
consolidation and/or recall, but not fear acquisition. This effect was driven by the B6 strain, as the D2 mice performed so poorly no additional effect of treatment was observed, as evidenced by a significant interaction between strain and treatment [interaction between strain and treatment F(1, 39) = 14.5, p < 0.001]. These effects were not due to non-specific effects on anxiety, activity, or overall growth as measured by open arm entries on the elevated plus maze (EPM), total arm entries, speed, and weight, respectively (Figure A-2).

Hp1bp3 knockdown alters specific aspects of the hippocampal transcriptome

We next wanted to determine the mechanism(s) by which hippocampal KD of Hp1bp3 may be impairing cognitive function. We performed total RNA sequencing on whole hippocampal lysates from 3 mice per strain/treatment group and confirmed KD of Hp1bp3 at the mRNA level (Figure 2-7A). A subset of genes was identified as differentially expressed relative to strain background, as well as relative to treatment group, while relatively few genes displayed a significant strain by treatment interaction, suggesting Hp1bp3 KD alters the hippocampal transcriptome comparably across background strain (Figure 2-7B and Supplemental Table 3). To identify specific pathways and functional categories altered by Hp1bp3 KD, we next utilized GSEA in order to identify GO terms significantly enriched among differentially expressed genes. Among genes up-regulated by Hp1bp3 KD, we identified a significant enrichment for immune-related terms, including response to interferon gamma, chemokine mediated signaling, and response to type I interferon (Figure 2-7C). This increase in inflammatory signaling seemed to be driven largely by an increase of microglia, as genes up-regulated by Hp1bp3 KD that also display a significant cell-type specific expression (Zhang et al., 2014) showed a significant overlap with a list of genes most highly expressed by microglia (hypergeometric test p < 0.001, Figure 2-7D). In contrast, among genes down-regulated by Hp1bp3 KD, we identified a significant enrichment for a number of terms related to neuron structure and function, including regulation of neuronal synaptic plasticity, excitatory synapse, and various channel activity terms (Figure 2-7E). In addition, among genes down-regulated by Hp1bp3 KD, there was a significant enrichment for transcripts experimentally observed to be most highly expressed in neurons and myelinating oligodendrocytes (hypergeometric test, p < 0.001).

As synaptic plasticity is a leading candidate mechanism for information and memory storage, and due to the fact that we identified regulation of ion channels and receptors as a candidate mechanism for Hp1bp3-induced cognitive impairment in an aging mouse population (Neuner et al., 2016), we were particularly interested in the most significantly down-regulated GO term, regulation of neuronal synaptic plasticity. The identification of this pathway was driven by 19 genes strongly down-regulated by Hp1bp3 KD (normalized enrichment score = -2.2, FDR = 0.01, Figure 2-8A). Of these 19 genes, 14 were identified as core driver genes by GSEA. Notably, the core driver gene most significantly down-regulated by treatment, particularly in B6 mice, was the activity-regulated cytoskeleton-associated protein (Arc), a gene repeatedly implicated in learning
Figure 2-7. Knockdown of *Hp1bp3* alters specific aspects of the hippocampal transcriptome.

(A) Knockdown of *Hp1bp3* was confirmed at the mRNA level [two-way ANOVA, effect of strain F(1, 8) = 2.6, p = 0.02, effect of treatment F(1,8) = 17.1, p = 0.003, no interaction between strain and treatment]. (B) *Hp1bp3* knockdown largely affects B6 and D2 mice similarly. Effect of strain: up-regulated genes are more highly expressed in B6 hippocampus. Effect of treatment: up-regulated genes are more highly expressed in *Hp1bp3* KD hippocampus. (C) Among genes up-regulated by *Hp1bp3* KD, there was a significant enrichment for immune-related GO terms. (D) There was a significant overlap between genes up-regulated by *Hp1bp3* KD that also display a cell-type specificity in their expression and genes most highly expressed in endothelial cells (EC) and microglia (hypergeometric test, p < 0.001). (E) Among genes down-regulated by *Hp1bp3* KD, there was a significant enrichment for terms related to neuronal excitability and function as well as synapse structure. (F) There was a significant overlap between genes down-regulated by *Hp1bp3* KD that also display a cell-type specificity in their expression and genes most highly expressed in neurons and myelinating oligodendrocytes (hypergeometric test, p < 0.001).
A. 

B. 

Effects of Strain

<table>
<thead>
<tr>
<th>Differentially Expressed Genes</th>
<th>Upregulated</th>
<th>Downregulated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ctrl</td>
<td>2620</td>
<td>1594</td>
</tr>
<tr>
<td>KD</td>
<td>1026</td>
<td>741</td>
</tr>
</tbody>
</table>

Effects of Knockdown

<table>
<thead>
<tr>
<th>Differentially Expressed Genes</th>
<th>Upregulated</th>
<th>Downregulated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ctrl</td>
<td>1642</td>
<td>901</td>
</tr>
<tr>
<td>KD</td>
<td>901</td>
<td>741</td>
</tr>
</tbody>
</table>

Interaction

<table>
<thead>
<tr>
<th>Differentially Expressed Genes</th>
<th>Upregulated</th>
<th>Downregulated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ctrl</td>
<td>12</td>
<td>3</td>
</tr>
<tr>
<td>KD</td>
<td>3</td>
<td>9</td>
</tr>
</tbody>
</table>

C. Positively Enriched GO terms

- Response to interferon gamma
- Chemokine mediated signaling
- Response to type I interferon
- Cytokine mediated signaling
- Chemokine receptor binding
- Response to virus
- Cellular response to interferon gamma
- Defense response to virus
- Regulation of symbiosis
- Neg. regulation of multi-organism process

Normalized Enrichment Score - log(FDR_{q})

D. 

E. Negatively Enriched GO terms

- Regulation of neuronal synaptic plasticity
- RNA splicing (transesterification)
- Neural nucleus development
- Excitatory synapse
- Cation channel activity
- Postsynaptic density
- RNP complex subunit org
- Regulation of synaptic plasticity
- VG cation channel activity
- VG K⁺ channel activity

Normalized Enrichment Score - log(FDR_{q})
Figure 2-8.  *Hplbp3* KD negatively regulates genes related to neuronal synaptic plasticity.

(A) Enrichment plot illustrating enrichment of the GO term regulation of neuronal synaptic plasticity among genes down-regulated by *Hplbp3* KD. Enrichment score = -0.6, Normalized enrichment score = -2.2, FDR = 0.01. (B) The core driver gene most significantly down-regulated by *Hplbp3* KD, particularly in B6 mice, was *Arc* (DESeq2 effect of treatment; log2FC = -1.2, adjusted p = 0.001, effect of strain; log2FC = -1.2, adjusted p = 0.001, interaction adjusted p = 0.2). (C) A number of plasma membrane proteins, particularly ion channels and receptors, were identified as additional core drivers of this enrichment. *DESeq2* main effect of treatment, p < 0.05.
and memory but for whom the upstream regulators remain poorly defined [(Shepherd and Bear, 2011), Figure 2-8B]. Additional core driver genes included a number of ion channels and receptors, including: 1) shisa family member 9 (Shisa9), an AMPAR auxiliary subunit (von Engelhardt et al., 2010), 2) junctophilin 3 (Jph3), a transmembrane junctional protein implicated in the regulation of the slow after-hyperpolarization (Moriguchi et al., 2006), 3) brevican (Bcan), an extracellular matrix protein that regulates the speed of synaptic transmission (Blosa et al., 2015), and 4) disks large homolog 4 (Dlg4), also known as post-synaptic density protein 95 (Psd95), which plays a critical role in organizing post-synaptic signaling (Chen et al., 2011). Functionally, these changes were shown to result in hippocampal neurons with decreased intrinsic neuronal excitability and decreased propensity to undergo LTP in slices from mice receiving Hplbp3 KD (data not shown here, see Neuner et al., 2018).

**Hplbp3 selectively regulates a subset of micro-RNAs**

Recent studies have demonstrated that HP1BP3 plays a role in micro-RNA (miRNA) biogenesis in human cells in vitro (Liu et al., 2016). To investigate whether observed transcriptional changes were due to a global change in miRNA biogenesis, we analyzed the small RNA sequences generated by our initial total RNA sequencing on hippocampal tissue from 3 mice/group. Of 1915 miRNAs identified and tested for differential expression analysis, only 35 showed differential expression relative to treatment at an adjusted p-value of 0.10 (Figure 2-9A and Supplemental Table 4), with only 9 remaining significant at an adjusted p = 0.05. Thus, less than 0.5% of all miRNAs tested exhibit significant differential expression following Hplbp3 KD. This is in comparison to mRNA, where just over 7% of all mRNAs tested were differentially expressed relative to Hplbp3 KD (proportions of differentially expressed genes significantly different, Z = 11.5, p < 0.001), suggesting Hplbp3 plays a much broader role in the regulation of mRNA expression over miRNA expression. These results demonstrate that in contrast to what has been observed in human cells (HeLa line), Hplbp3 does not seem to be a global regulator of miRNA biogenesis in vivo in the mouse hippocampus.

Although global changes in miRNA expression were not observed, we next tested the hypothesis that the select miRNAs regulated by Hplbp3 may explain downstream observed changes in mRNA expression and synaptic function. First, we input the list of mRNAs differentially expressed relative to treatment into IPA software (Kramer et al., 2014) and used the Core Analysis function to identify putative miRNA upstream regulators of our observed mRNA changes (Table A-1). A number of miRNAs were identified as putative upstream regulators, as measured by significant overlap of known target genes and genes present in our list of differentially expressed mRNA. The top predicted upstream regulator, mir-21, was predicted to be inhibited in our dataset based on the direction of change of its target molecules. However, the pattern of mir-21 differential expression after Hplbp3 KD did not fit this pattern, eliminating mir-21 as a miRNA that could possible explain mRNA changes observed in our hands (Table A-1).
Figure 2-9. *Hp1bp3* selectively regulates the miRNA transcriptome in vivo. (A) Plot illustrating (x-axis) average miRNA expression versus (y-axis) log Fold Change as measured by DESeq2. Red dots illustrate those miRNA transcripts that were determined to be differentially expressed relative to treatment at an adjusted p = 0.10. (B) miRNA mmu-mir-223-3p was identified as a putative upstream regulator of observed mRNA changes after *Hp1bp3* KD and was predicted to be activated in our dataset by IPA. mmu-mir-223-3p was significantly up-regulated by *Hp1bp3* knockdown [*two-way ANOVA, effect of treatment F(1, 8) = 5.6, p = 0.046, no significant effect of strain or interaction between strain and treatment*].
We next examined the miRNA identified as the second most likely miRNA upstream regulator, mir-223. Based on the direction of change of its target molecules, IPA predicted that mir-223 was activated in our dataset and indeed, mir-223 (specifically mmu-mir-223-3p) was observed to be up-regulated after Hplbp3 KD (two-way ANOVA, effect of treatment F(1, 8) = 5.6, p = 0.046, Figure 2-9B) and was one of the miRNA transcripts identified as significantly differentially expressed by DESeq2 at an adjusted p = 0.10 (mmu-mir-223-3p adjusted p = 0.09, log2FC = 1.2). mir-223, best studied in the context of hematopoietic stem cells, has been shown to function as a key modulator in the differentiation and activation of myeloid cells and has been implicated in a number of immune system functions (Yuan et al., 2018). In accordance with these known functions, the list of mir-223 targets differentially expressed after Hplbp3 KD was identified to be most highly enriched for the GO Biological Process term “immune system process” (FDR < 0.001) through WebGestalt over-representation enrichment analysis (Wang et al., 2013). Notably, mir-223 has been shown to promote maturation, proliferation, and activation of myeloid cells (Tsitsiou and Lindsay, 2009), which overlaps with our previous observation that the upregulation of immune-related genes following Hplbp3 KD is likely due to an increase in numbers of microglia (Figure 2-7D).

Transcriptome-level changes induced by Hplbp3 knockdown recapitulate those observed in human aging

In addition to standard GO terms and functional pathways, GSEA also allows gene lists of interest to be compared against experimentally derived gene sets that have been uploaded into MSigDb (Liberzon et al., 2011). When we compared our list of differentially expressed mRNAs to the chemical and genetic perturbations (CPG) database, we observed significant overlap between genes down-regulated by Hplbp3 KD and those genes down-regulated with age in the human frontal cortex [(Lu et al., 2004), FDR q-value = 0.04, Figure 2-10A]. This significant enrichment was driven by 43 genes significantly down-regulated in both the aging human frontal cortex and in the mouse hippocampus following Hplbp3 KD, 37 of which were determined by GSEA to be core drivers of enrichment. The twenty most significantly down-regulated core drivers, as well as their corresponding fold changes from the human brain (Lu et al., 2004) are listed in Figure 2-10B. As a group, this set of commonly down-regulated genes displayed relatively diverse enrichment for GO Biological Process terms using WebGestalt over-representation enrichment analysis (Wang et al., 2013), including interleukin secretion (IL-5 FDR = 0.03, IL-13 FDR = 0.03), microtubule polymerization (FDR = 0.04), and chemical synaptic transmission (FDR = 0.04), all processes known to be altered in aging. In summary, the overlap of the broader Hplbp3 KD transcriptome with that of human aging supports the idea that Hplbp3 is a critical regulator of a number of the transcriptional changes that occur during aging and may serve as a valuable therapeutic target to delay or prevent aging-related cognitive decline.
Figure 2-10. Transcriptome-level changes induced by *Hp1bp3* knockdown recapitulate those observed in human aging.

(A) Enrichment plot illustrating overlap of genes down-regulated in the human aging cortex with genes down-regulated by *Hp1bp3* KD. Enrichment score = -0.4, Normalized enrichment score = -1.9, FDR = 0.04. (B) List of twenty genes down-regulated by *Hp1bp3* KD and their corresponding fold changes in the human aging cortex as reported by Lu et al. (2004).
HP1BP3 is associated with memory function in human aging

To test the hypothesis that *Hp1bp3* itself is misregulated in human cognitive aging and to evaluate the translational relevance of our mouse results, we next tested whether hippocampal *Hp1bp3* expression correlated with cognitive performance in elderly humans (Table 2-2). Caveats to working with human brain samples include variable post-mortem intervals and undefined brain pathology (Bennett et al., 2014; Hargis and Blalock, 2017). Therefore, in order to limit potential confounds we analyzed hippocampal tissue from humans with ‘normal’ cognitive aging (i.e. no distinct Alzheimer’s disease pathology), which is not typically associated with gross neurodegeneration (Burke and Barnes, 2006; Korbo et al., 2004). Quantitative Western blots for HP1BP3 were performed on postmortem hippocampal tissue lysates prepared from cognitively intact [mini-mental status exam (MMSE) = 29.3 ± 0.3, age = 78 ± 9.7 years] and cognitively impaired humans (MMSE = 22.6 ± 3.0, age = 84 ± 4.4 years). HP1BP3 levels were significantly lower in humans diagnosed with cognitive impairment compared to age-matched controls (Figure 2-11). Double band staining is typical of HP1BP3 expression, indicative of multiple splice variants present in brain tissue (Garfinkel et al., 2015b). These results suggest either lower or reduced expression of *Hp1bp3* may underlie cognitive deficits observed in human populations, and highlight the value and potential relevance of aged cohorts of diverse strains of mice for identifying molecular correlates of cognitive decline in humans.

Discussion

The goal of this study was to utilize a genetically diverse panel of mice in order to identify genetic factors involved in the regulation of cognitive aging that may have gone undetected in either complex human studies or murine studies utilizing only a single genetic background. Aging is a leading risk factor for age-associated dementias such as Alzheimer’s disease, and our work and others suggest that genetic factors and mechanisms underlying biological processes during midlife play a key role in determining an individual’s susceptibility or resilience to transitioning between healthy brain aging and pathological brain aging (Douaud et al., 2013; Jack et al., 2013; Kaczorowski et al., 2011; Keller, 2006; Miller et al., 2008). Thus, there is a critical need to understand gene variants that play a role in determining memory capabilities at this transition point (midlife). To directly address this need, and to overcome some of the barriers inherent to human studies, we turned to a well-characterized mouse GRP, which allows for the exploitation of phenotypic heterogeneity across a population while exerting precise control of environmental conditions.

Identification and validation of *Hp1bp3* as a modifier of cognitive decline

By combining forward and reverse murine genetic approaches and by joint analysis of mouse and human cohorts, we progressed from the identification of a
Table 2-2.  Demographics for elderly humans used for postmortem hippocampal tissue analysis.

<table>
<thead>
<tr>
<th>Patient</th>
<th>Age</th>
<th>Sex</th>
<th>MMSE</th>
<th>Group</th>
<th>Post-Mortem Interval (PMI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>79</td>
<td>Female</td>
<td>29</td>
<td>Intact</td>
<td>1.75</td>
</tr>
<tr>
<td>H2</td>
<td>93</td>
<td>Female</td>
<td>30</td>
<td>Intact</td>
<td>2.25</td>
</tr>
<tr>
<td>H3</td>
<td>81</td>
<td>Male</td>
<td>29</td>
<td>Intact</td>
<td>2.83</td>
</tr>
<tr>
<td>H4</td>
<td>86</td>
<td>Female</td>
<td>27</td>
<td>Impaired</td>
<td>2.48</td>
</tr>
<tr>
<td>H5</td>
<td>90</td>
<td>Female</td>
<td>24</td>
<td>Impaired</td>
<td>2.50</td>
</tr>
<tr>
<td>H6</td>
<td>59</td>
<td>Male</td>
<td>17</td>
<td>Impaired</td>
<td>5.85</td>
</tr>
</tbody>
</table>

MMSE; Mini-Mental State Exam score (maximum score indicating normal cognition = 30)

Figure 2-11.  HP1BP3 protein is correlated with memory status in aging humans.  
A) Western blot shows hippocampal HP1BP3 protein is reduced in cognitively impaired elderly humans (n = 3, MMSE = 22.6 ± 3.0, age = 84 ± 4 years) relative to cognitively intact controls (n = 3, MMSE = 29.3 ± 0.3, age = 78 ± 10 years).  B) Quantification of hippocampal HP1BP3 protein (Cog. Intact = 1.02 ± 0.21 adjusted density, Impaired = 0.60 ± 0.04, t(1,4) = -3.344, p = 0.03).
significant QTL containing variants regulating CFM at midlife to the demonstration that our top positional candidate, *Hp1bp3*, is important for hippocampus-dependent long-term and spatial working memory. Specifically, we demonstrate for the first time that a targeted KD of *Hp1bp3* in the hippocampus by 50-75% was sufficient to recapitulate cognitive deficits observed in global *Hp1bp3* KO mice. No effect on CFA was observed, implicates a relatively specific role for *Hp1bp3* in mechanisms underlying working memory as well as memory consolidation and recall, but highlights that CFA may rely on independent mechanisms for successful completion. Overall, results here demonstrate that cognitive deficits caused by *Hp1bp3* deficiency, presumably in both KO mice and in aged impaired mice, are due primarily to *Hp1bp3* deficiency in the hippocampus. This is an important finding, as *Hp1bp3* KO mice are smaller than WT littermates and have alterations in insulin signaling that may have impacted cognitive performance due to peripheral mechanisms (Garfinkel et al., 2015a). In addition, we extend our previous finding that loss of *Hp1bp3* has a negative impact in cognitive function and provide mechanistic insight into how these alterations occur.

Specifically, we show *Hp1bp3* KD negatively regulates a number of genes critical for synaptic plasticity and neuronal activity, including the immediate early gene *Arc* (Figure 2-8B). Although *Arc* is well known to be critical for memory formation (Guzowski et al., 2000; Plath et al., 2006), the precise signaling cascades involved in *Arc* transcription and regulation are not well defined (Shepherd and Bear, 2011). Here, we discover a novel role for *Hp1bp3* as a regulator of *Arc* expression, which implicates *Hp1bp3* in the broader context as a master regulator of transcriptional network changes required for successful induction of synaptic plasticity, likely via its role in modifying chromatin structure (Dutta et al., 2014). Finally, we demonstrate that observed transcriptional changes functionally manifest as impairments in both intrinsic neuronal excitability as well as synaptic plasticity (data not shown here, see Neuner et al., 2018). Together, these results provide important insight into physiological functions of *Hp1bp3* under baseline conditions in adult mice. As a result, we can now better understand how misregulation of *Hp1bp3* may contribute to cognitive impairments seen in both *Hp1bp3* KO mice and cognitively impaired aging mice and humans with decreased levels of *Hp1bp3* (Neuner et al., 2016).

**Hp1bp3** effect on cognition is robust to genetic context

It has long been known that genetic background is critical for modifying phenotypic presentation. Recently, this was highlighted by a survey of 30 inbred lines, where in some cases, opposite conclusions were drawn regarding the effect of gene KO depending on the genetic context of the manipulation (Sittig et al., 2016). Therefore, we thought it important to evaluate the effect of *Hp1bp3* manipulation in more than one genetic background. We performed our viral experiments in the B6 and D2 inbred strains, as these strains were used to derive the BXD GRP that was originally used for identifying *Hp1bp3* as a modifier of cognitive aging. These two strains show marked differences in learning and memory abilities (Neuner et al., 2016; Philip et al., 2010) as well as differences in certain types of LTP (Matsuyama et al., 1997; Nguyen et al., 2000). In
general, our data are consistent with this literature, with the B6 strain outperforming the D2 strain on CFC. Certain types of LTP have been shown to be reduced in the D2 strain (Schimanski and Nguyen, 2005), which may explain their poor performance on CFC—a task which requires synaptic plasticity (Tang et al., 1999). Poor performance on CFC may also be explained, in part, by lower baseline levels of \(H_{p1bp3}\) mRNA observed in the D2 strain (Figure 2-7), which while not detected at the protein level (Figure 2-6), is consistent with our findings that BXD strains carrying the \(D\) allele of \(H_{p1bp3}\) show decrease hippocampal \(H_{p1bp3}\) expression (Figure 2-1). As \(H_{p1bp3}\) KD impaired cognitive function in both strains, as well as largely affected the transcriptome across strains similarly (Figure 2-7), our data suggests \(H_{p1bp3}\) deficiency has a robust detrimental effect on cognitive function regardless of genetic context. This broadens the translational relevance of this finding, as restoring \(H_{p1bp3}\) levels (or supplementing with a small molecule that increases \(H_{p1bp3}\) levels) is likely to be beneficial to a wider variety of subjects than a target with context-dependent effects.

**\(H_{p1bp3}\) as a regulator of miRNA biogenesis**

Contrary to previous reports (Liu et al., 2016), \(H_{p1bp3}\) does not seem to be a global regulator of miRNA biogenesis, at least in the adult mouse hippocampus. Multiple explanations for these conflicting results exist, including different experimental species (mouse versus human), biological source (mouse CNS versus human cervical cancer), differences in experimental techniques (in vitro versus in vivo), and even differences in miRNA alignment protocols. In addition, we did observe a select set of miRNAs differentially expressed relative to \(H_{p1bp3}\) KD, suggesting \(H_{p1bp3}\) does regulate biogenesis of at least some miRNAs in vivo. Of particular interest to this study is mir-223, which was identified by IPA as a putative upstream regulator of observed mRNA-level changes following \(H_{p1bp3}\) KD and whose expression in our miRNA sequencing data changed concordantly with that predicted by IPA. As a number of immune-related genes have been shown to be down-regulated in mir-223 deficient cells (Lu et al., 2013), the large number of immune-related genes up-regulated after \(H_{p1bp3}\) KD may, in part, be due to observed up-regulation of mmu-mir-223-3p. While mir-223 is best studied in the context of the hematopoietic system, recent studies have investigated its role in CNS functions and have identified enriched expression of mir-223 in the hippocampus relative to other brain regions (Harraz et al., 2012). In particular, mir-223 has been shown to decrease total dendritic tree length, branch number, and complexity in cell culture (Harraz et al., 2014) and specifically target glutamate receptor transcripts for degradation in vivo (Harraz et al., 2012), which may link mir-223 to some of the additional neuronal phenotypes observed in \(H_{p1bp3}\) KD mice. In summary, while miRNA biogenesis likely contributes to some of the transcriptional changes observed after \(H_{p1bp3}\) KD, additional mechanisms regulating gene expression are likely at play, including alterations in chromatin structure that result in changes in the accessibility of key genomic regions to necessary transcriptional machinery.
**Hp1bp3 knockdown recapitulates symptoms of aging and Alzheimer’s disease**

Many alterations observed after Hp1bp3 KD appear to phenocopy alterations observed in aging and AD (Neuner et al., 2018). In particular, studies have identified decreases in neuronal excitability as cellular correlates of cognitive impairments in aging animals (Disterhoft et al., 2004). Notably, the slow AHP is increased in aged impaired animals relative to young and aged non-impaired animals (Kaczorowski and Disterhoft, 2009). Similar neuronal phenotypes have been observed in animal models of AD (Kaczorowski et al., 2011). As these aged and AD neurons are less able to respond adequately to input, this reduced neuronal excitability translates into a reduction in the maintenance of LTP and poorer memory storage. Transcriptionally, mRNA changes after Hp1bp3 KD significantly overlap mRNA changes observed in the aging human cortex [FDR < 0.05, **Figure 2-10**, (Lu et al., 2004)] and are reminiscent of increased inflammation and neurodegeneration observed in AD patients (Van Eldik et al., 2016). In addition, baseline reductions in Arc expression have been observed in aging rodents, which correlates to less efficient memory storage and retrieval during aging (Penner et al., 2011). Each of these hallmark features of aging and AD - decreased neuronal excitability, reduced synaptic plasticity, impaired cognitive function, and increased neuroinflammation – are recapitulated in our Hp1bp3 KD animals. Together with our previous observations that cognitively impaired mice and humans have naturally occurring reductions in Hp1bp3 levels, our data suggest reduced Hp1bp3 is a relevant driver of aging and AD-related phenotypes. As aging is the leading risk factor for many diseases in addition to AD, and there is currently no cure for AD, understanding the origin of these deficits is critical for developing effect therapeutics to maintain healthspan in aging individuals. The identification and characterization of Hp1bp3’s effect on cognition and neural function under baseline adult conditions is a critical first step toward understanding how aging, Hp1bp3, and additional genetic and environmental factors interact to modify an individual’s susceptibility to age-related cognitive decline.

**Hp1bp3 may play conserved role in humans**

Reduced expression of Hp1bp3 is observed in both cognitively impaired aged mice and humans, suggesting that decreased expression of Hp1bp3 contributes to cognitive aging in both species. In support of this idea, Hp1bp3 is among the top 100 genes up-regulated in response to metformin hydrochloride (Lamb et al., 2006), a drug known to enhance memory in mice (Wang et al., 2012), which was recently approved for clinical trial to prevent or reduce effects of aging, including cognitive decline (Targeting Aging with Metformin, TAME study). In addition, Hp1bp3 is highly conserved in mammals, with 93% similarity between the human and murine primary sequences (Garfinkel et al., 2015b) indicating Hp1bp3 likely plays an important functional role in both species. Deletion of the region of the human genome syntenic to our QTL results in a condition known as Deletion 1p36 syndrome (Battaglia et al., 2008). Although specific phenotypes vary according to size and location of deletion breakpoint (Gajecka et al., 2007), many with this syndrome exhibit cognitive deficits and mental retardation.
(Battaglia et al., 2008), supporting the idea that genes in this area are necessary for cognitive function in both humans and animals.
CHAPTER 3. HARNESSING GENETIC COMPLEXITY TO ENHANCE MOUSE MODELS OF AD: A PATH TOWARD PRECISION MEDICINE*

Introduction

AD is a neurodegenerative disorder characterized by both dementia and the accumulation of neuropathological amyloid plaques and tau tangles (Selkoe, 1991). Mutations that drive overproduction of beta-amyloid (Aβ) have been shown to cause early onset FAD, leading to a model in which production and accumulation of Aβ is thought to be an initiating event in a sequence leading to memory loss, neurodegeneration, gliosis, and synaptic dysfunction (Hardy and Higgins, 1992). However, strategies to directly target amyloid for clearance have failed to translate into successful treatments and the number of deaths attributable to AD and costs associated with the disease continue to rise.

Although age is the greatest risk factor for developing AD, it is increasingly clear that genetics and family history play a large role. The heritability of AD is estimated to be in the range of 50-80%, indicating an individual’s susceptibility or resilience to disease is, at least in part, determined by heritable DNA variants (Gatz et al., 1997). Even among patients with FAD mutations, the age at first symptom onset is widely variable, with some patients exhibiting symptoms decades later than predicted based on mutation status. This suggests additional genetic factors exist that may provide protection from disease (Ryman et al., 2014). While a number of genetic risk factors have been identified (Lambert et al., 2013), besides APOE and TREM2, effect sizes of identified variants are generally small. As such, a large proportion of variation in disease risk and severity remains undefined and unexploited (Ridge et al., 2013). In addition, the study design utilized in most genome wide association tests (e.g. case-control) is primed to identify variants associated with risk of AD, rather than variants that modify an individual’s disease trajectory and/or delay the onset of disease and provide protection from AD. This is due primarily to the fact that asymptomatic individuals rarely enter the clinic for treatment, and even if they are included in a study, are likely to be enrolled as cognitively normal controls.

Identifying genetic variants and pathways involved in protection from AD will provide valuable targets for new therapeutics to prevent or delay the onset of symptoms. Individuals with high-risk genotypes but who fail to present with clinical symptoms of AD represent an ideal population in which to study resilience. However, causal mutations

in APP and PSEN1/2 (i.e. high-risk genotypes) are rare in humans, greatly limiting statistical power and opportunity for analysis. In addition, access to brain tissue at early disease time points, before overt symptom onset, is limited for obvious reasons in human studies, precluding the identification of causal versus collateral molecular mechanisms. Thus, mouse models harboring causal AD mutations are important tools that present many advantages including defined high-risk genotypes, early access to brain tissue, and precise environmental control.

That said, there are caveats to traditional AD mouse models, including the fact that most mouse models of AD are maintained on only a single or a few genetic backgrounds (Onos et al., 2016). This includes backgrounds of mixed origins, which may confound interpretation of results. Previous studies in mice have demonstrated that genetic background has a strong effect on Aβ levels, with several mapping studies in F2 populations identifying large genomic regions involved in modulating amyloid burden (Ryman et al., 2008; Sebastiani et al., 2006). However, these studies and individual strain-by-strain comparisons have not evaluated the effect of genetic background on cognitive performance (Jackson et al., 2015; Sipe et al., 1993). Given the impact of genetics on cognitive decline in human populations (Gatz et al., 1997), we hypothesized genetic background was also playing an important role in modifying cognitive decline in animal models of disease, and that inclusion of genetic diversity would improve translational validity of AD mouse models.

In order to directly test this hypothesis, and to aid in the identification of specific genes involved in modifying resilience to AD, we developed the first AD transgenic mouse GRP. This panel, which we term the AD-BXDs, combines two well-established resources: 1) the 5XFAD transgenic line on an otherwise fully inbred B6 background that recapitulates various aspects of the human disease, including Aβ42 accumulation, cognitive deficits, and neuron loss (Oakley et al., 2006), and 2) the BXD GRP, the largest and best-characterized series of recombinant inbred strains derived from the two common inbred strains B6 and D2 (Peirce et al., 2004; Taylor et al., 1999). The BXD panel segregates for more than 4.8 million sequence variants, including many in genes known to confer risk for AD (Wang et al., 2016). The resulting panel of F1 hybrids represent a novel and fully isogenic resource to monitor phenotypic outcomes in individuals harboring identical high-risk FAD mutations in human APP and PSEN1 genes, raised in controlled environments, but whose allelic contributions differ across the remainder of the genome.

The general aim of the work described here is to build and test the validity of the AD-BXDs as a resource that will enable the research community to systematically identify sets of genetic variants and pathways involved in determining individual susceptibility or resilience to AD. As both parental lines of the AD-BXD panel are fully inbred, the resulting panel also provides a reproducible resource to efficiently evaluate gene-by-environment-by-treatment effects to test, triage, and translate therapeutics more quickly and accurately. To evaluate whether this design produced a valid model of AD, here we show that our AD-BXD panel faithfully recapitulates key aspects of the human disease, including phenotypic variation in disease onset and severity, sensitivity to
genetic variation in genes known to confer risk for human LOAD, and a high level of concordance with transcriptional aspects of human disease. Thus, we present the AD-BXD panel as a new mouse model of human AD with high translational potential for both understanding the complex etiology of FAD and sporadic LOAD and discovering new genetic and molecular pathways associated with AD risk and resilience.

**Methods**

**Animals**

Female congenic B6 mice hemizygous for the dominant 5XFAD transgene (Oakley et al., 2006), which consists of 5 human mutations known to cause familial AD [three in amyloid precursor protein (*APP*; Swedish: K670N, M671L, Florida: I716V, and London: V717I) and two in presenilin 1 (*PSEN1*; M146L and L286V)], were obtained from The Jackson Laboratory (JAX MMRRC Stock No: 34848-JAX). These mice were bred with 28 males from a set of genetically diverse recombinant inbred strains from the well-established BXD GRP (Peirce et al., 2004). By selecting the same maternal background strain (i.e. 5XFAD-B6) across the panel for cross with male BXD strains, we were able to introduce variants in the nuclear DNA, hold the mitochondrial genome constant, and control for strain-specific differences in maternal behavior on offspring behavior. The F1 progeny resulting from this B6-5XFAD by BXD cross are isogenic recombinant inbred backcross mice, each harboring one maternally derived B allele and either a B or D paternally derived allele at any given genomic locus. As expected from a Mendelian pattern of inheritance, ~50% of these F1 mice carry the 5XFAD transgene (termed AD-BXDs) and ~50% are non-transgenic littermate controls referred to Ntg-BXD. Male and female offspring were group housed (2-5 per cage) and maintained on a 12 hr. light/dark cycle with ad libitum access to food and water. All mice were genotyped for the 5XFAD transgene through a combination of in-house genotyping according to The Jackson Laboratory protocols for strain #34848-JAX and outside services (Transnetyx, TN, USA, and The Jackson Laboratory Transgenic Genotyping Services). Working memory and body weights were monitored longitudinally, and more detailed phenotyping occurred at 6 and 14 months. These time points were selected to obtain an adult phenotype (6 months) and a middle-aged to aged time point (14 months) that captured variation in disease symptoms before the mice exhibited severe health-related problems that may confound behavioral testing. All mouse experiments occurred at University of Tennessee Health Science Center and were carried out in accordance with the standards of the Association for the Assessment and Accreditation of Laboratory Animal Care (AAALAC), as well as the recommendations of the National Institutes of Health Guide for the Care and Use of Laboratory Animals. The Institutional Animal Care and Use Committee (IACUC) at the University of Tennessee Health Science Center approved the protocol.
Behavioral assays

Y-maze

For all behavioral testing, mice were habituated to transport and to the testing room for three days prior to testing. The y-maze test of spontaneous alternation was performed as described previously (Oakley et al., 2006). The y-maze used for testing was made of clear acrylic with arms that were 2” wide x 12” long x 2” high. The maze was placed on a table in a dimly lit room and spatial cues were displayed on walls around the table. Mice were placed in a randomized start arm and video tracking software was used to monitor arm entries (ANY-maze, Stoelting Co., IL, USA). An arm entry was called when the mouse’s entire body, including the two back feet, entered the arm. The sequence and total number of arms entered was recorded, and the percentage of successful alternations was calculated as follows: number of alternations/maximum possible alternations (total number of arms entered – 2) x 100. For each animal that was measured longitudinally (i.e. not harvested at the early 6 month time point), the age at which each animal became ‘impaired’, or performed below chance levels (50%), was recorded and used as the animals “age at onset” [AD-BXDs: n = 226 (126 females/100 males) across 28 strains vs Ntg-BXDs, n = 171 mice (108 females/63 males) across 25 strains]. Strain averages for age at onset were then calculated.

Sensorimotor battery

At 6 months [AD-BXDs n = 284 (185 females/90 males) across 28 strains, Ntg-BXDs n = 220 (158 females/62 males) across 27 strains] and 14 months [AD-BXDs n = 222 (104 females/106 males) across 26 strains, Ntg-BXDs n = 172 (109 females/63 males) across 25 strains], mice were subjected to a sensorimotor battery consisting of three tasks. First, mice were placed in the center of a 3-foot long narrow (0.5”) beam elevated 20.75” off a table surface and the time taken for the mouse to cross the narrow beam onto a safe platform on either side was measured. Second, mice were placed face-down on a wire mesh grid (holes were 1cm x 1cm) that was placed at a 45° angle. The time taken for a mouse to right itself (negative geotaxis) was recorded. A 3 minute maximum time limit was imposed for both the narrow beam and incline screen tests. If a mouse fell from the narrow beam, the maximum score of 180s was given. Third, grip strength was measured using a standard grip strength meter (Coulbourne Instruments). Each of these three tasks were repeated in triplicate and the average score across three trials was used. For each mouse, a z-score based on the 6 month-old population average was calculated for each task and the three z-scores were summed to derive a sensorimotor composite score, which was used here to relate sensorimotor performance to cognitive abilities.

Elevated plus maze

At 6 months [AD-BXDs n = 280 (191 females/89 males) across 28 strains, Ntg-BXDs n = 220 (158 females/62 males) across 27 strains] and 14 months [AD-BXDs n = 221 (116 females/105 males) across 26 strains, Ntg-BXDs n = 173 (110 females/63
males) across 25 strains], anxiety was evaluated using an EPM task. Mice were placed in the center of the maze and allowed to explore for 6 minutes. Video tracking software (ANY-maze, Stoelting Co.) was used to track the mouse and calculate the time spent in open versus closed arms of the maze as well as the number of arm entries into either open or closed arms, the total number of arm entries, and the total distance travelled in the maze.

**Contextual fear conditioning**

Following three days of habituation to transport and to the testing room, mice were trained on a standard CFC paradigm as previously described (Neuner et al., 2015). Training consisted of a 180s baseline period followed by four mild foot shocks (1s, 0.9mA), separated by 115 ± 20s. A 40s interval following each foot shock was defined as the post-shock interval, and the percentage of time spent freezing during each of these intervals was measured using FreezeFrame software (Coulbourne Instruments, PA, USA). The percentage of time spent freezing during the final post-shock interval (PS4) was used as an index of CFA. Twenty-four hours later, hippocampus-dependent CFM was tested by returning the mouse to the testing chamber for 10 min. The percentage of time spent freezing during the testing trial was measured using FreezeFrame software and used as an index of CFM. For CFC, 146 6 month AD-BXD (102 females/44 males) and 209 14 month AD-BXD (111 females/98 males) across 26 strains were used, along with 114 6 month Ntg-BXD (83 females/31 males) across 24 strains and 167 14 month Ntg-BXD mice (106 females/61 males) across 27 strains. Pain sensitivity was evaluated in a subset of mice by recording the length of activity burst following each shock. An average post-shock reactivity score was calculated by averaging the length of each activity burst following the four training shocks.

**Heritability estimates**

Heritability estimates for each phenotype were calculated according to established methods (Belknap, 1998). Briefly, we compared between-strain variance (due to genetic diversity, $V_G$) to total sample variance (due to both genetic and environmental factors, $V_E$) given the average number of biological replicates per strain ($n$) according to the following formula (Equation 3-1):

$$h^2_{RX} = \frac{V_G}{V_G + V_E/n}$$  \hspace{1cm} (Eq. 3-1)

The average number of mice per strain was used to represent $n$. $V_E$ was calculated by summing between-strain variance and within-strain variance, as within-strain variance should capture all variation not due to genetic diversity. As heritability was calculated using both males and females, within-strain variance will also capture variation due to sex. However, as we calculated heritability independently for each trait of interest across AD- and Ntg-BXDs, our heritability estimates do not capture variation due to age or genotype.
Enzyme-linked immunosorbent assay (ELISA)

Brains were removed immediately following CFM testing at appropriate time points (6 months or 14 months) and hemisected. One half of the brain was immediately dissected, snap frozen, and stored at -80°C until use. Aβ42 levels were quantified from sections of temporal cortex [6 month n = 72 mice (46 female/46 male) across 22 AD-BXD strains, 14 month n = 82 mice (43 female/33 male) across 21 AD-BXD strains] as previously described (Oakley et al., 2006). Briefly, tissue was homogenized in 1X PBS + 1% Triton-X 100 using the TissueLyser II system (Qiagen) and sonicated 2x10s on low power. Protein concentration was determined using a NanoDrop 2000 UV-Vis Spectrophotometer (ThermoScientific, USA). Brain homogenates (10 mg/ml) were extracted in a final concentration of 5M GuHCl overnight at 4°C. Samples were then diluted appropriately and run in duplicate on Aβ42-specific sandwich colorimetric ELISAs according to the manufacturer’s protocol (Cat# 298-92401, Wako Chemicals, Richmond, VA). Optical densities at 450 nm were read on a Biotek plate reader (BioTek, USA) and Aβ42 concentration was determined by comparison with Aβ42 standard curves. Only readings in the linear range of the standard curve were included in analysis. Duplicates were averaged to determine concentration of Aβ42 in each sample. Finally, Aβ42 concentrations were normalized to total protein concentration and are reported as nanograms of Aβ42 per milligrams of total protein.

Immunohistochemistry and plaque quantification

At harvest, the half brain not used for fresh dissection was placed in 4% paraformaldehyde and kept at 4°C until further use. In order to minimize technical variation in immunohistochemistry (IHC), brains were sent to Neuroscience Associates (Knoxville, TN), where 40 hemibrains were embedded, processed, and stained simultaneously. Briefly, the brains were freeze-sectioned coronally at 40 μm intervals (not including cerebellum) and staining for Aβ42 was performed on every 24th section spaced at 960 μm, yielding approximately 9 sections per hemibrain. For analysis, images of each section containing hippocampus were collected on a Nikon Eclipse 90i microscope using NIS-Elements Advanced Research program. Images were taken using a 2x objective with computer automated focusing. Approximately 4 images were captured for each hemibrain and stitched together using NIS-Elements Advanced Research program. ImageJ particle analysis was used to automate detection of plaques (Hurtado et al., 2010). Regions of interest (hippocampus and cortex) in each image were manually outlined and pixel size of each region calculated and used to determine the percentage of each area covered by amyloid plaques, controlling for regional size differences. Neuroscience Associates also performed scanning of each slide at 20x using a Huron scanner and these images are used for illustrative purposes.
Transcriptome analyses

RNA sequencing

Snap frozen hippocampi from AD-BXD strains and Ntg-BXD littermate controls at 6 month [AD-BXD n = 33 (15 females/18 males) across 13 strains, Ntg-BXD n = 31 (17 females/14 males) across 14 strains] and 14 months [AD-BXD n = 36 (16 female/20 male) across 14 strains, Ntg-BXD n = 33 (17 female/16 male) across 15 strains] were used for RNA sequencing. RNA was isolated on a Qiacube using the RNeasy mini kit (Qiagen) and treated with DNase to remove contaminating DNA. RNA quality was confirmed using a BioAnalyzer (Agilent Technologies). All samples had RNA Integrity Numbers (RIN values) > 8.0. Sequencing libraries were prepared from 1 μg RNA with the Truseq Stranded mRNA Sample Preparation Kit (Illumina Inc) following the manufacturer’s protocol. Final PCR-enriched fragments were validated on a 2200 Tapestation Instrument using the D1000 ScreenTape (Agilent Technologies) and quantified by qPCR using a Universal Library Quantification Kit (Kapa Biosystems) on the QuantStudio 6 Flex (ThermoFisher Scientific). Final library pools were sequenced by 75bp paired-end sequencing on a HiSeq2500 (Illumina Inc). Because both B6 and D2 alleles segregate within our panel, the GPRS pipeline developed by the Churchill group at The Jackson Laboratory was used (https://github.com/churchill-lab/gbrs; Choi et al., 2017). This pipeline utilizes Bowtie to first align reads to diploid B6D2 transcriptome, followed by an expectation maximization algorithm in order to quantify the number of reads aligned to a given allele (Raghupathy et al., 2018). This allows for the quantification of both total reads assigned to a gene and the number of reads assigned to either the B or D allele. For final by-strain analysis, samples belonging to the same strain/sex/age/genotype group were averaged. Differential expression analysis was conducted using the DESeq2 package (Love et al., 2014). For evaluation of transgene expression and its effect on endogenous *App* and *Psen1* expression, RNA sequencing reads from a larger subset of AD- and Ntg-BXD [n = 293 (177 females/116 males across 28 strains)] were sequenced according to identical methods and were additionally aligned to the mutated human *APP* and *PSEN1* sequences. Expression was quantified using transcripts per million and then log transformed to compare expression across groups.

Comparison of AD-BXD and human transcriptomes

In order to evaluate how well the AD-BXD transcript profile matches that of human AD, we utilized a dataset recently published by Hargis and Blalock (Hargis and Blalock, 2017) comparing existing mouse models of AD to human AD. They identified a consensus AD signature consisting of 60 genes derived from the top 10% commonly up-regulated and down-regulated genes across three human AD datasets (Supplemental Table 5). In order to see how the transcriptome from our AD-BXD panel compared to normal expression patterns, differential expression analysis comparing hippocampal gene expression from 14 month-old AD-BXD lines to non-carrier littermate controls was performed using DESeq2 (Love et al., 2014). The log$_2$ fold change (log$_2$FC) for each of the 60 AD consensus genes that were significantly differentially expressed (nominal p-value < 0.05) across AD and Ntg-BXD was identified and used for comparison across
human and mouse datasets obtained from Hargis et. al. (Hargis and Blalock, 2017). To evaluate similarities between immune-enriched genes up-regulated in AD-BXDs and gene lists identified as associated with AD from Zhang et. al. (Zhang et al., 2013) and Jones et. al. (International Genomics of Alzheimer's Disease, 2015), all three gene lists were uploaded into GeneWeaver (Baker et al., 2016) and Jaccardian similarity indexes were calculated and evaluated for significance.

**Gene set enrichment analysis**

GSEA was performed as previously described (Neuner et al., 2016; Subramanian et al., 2005). Briefly, significantly differentially expressed genes from the comparison of interest (**Supplemental Table 6**: All AD-BXD vs All Ntg-BXD, **Supplemental Table 7**: 14 month AD-BXD vs 14 month Ntg-BXD and 6 month Ntg-BXD vs 14 month Ntg-BXD) were ranked according to log2 fold change. GO gene sets were obtained from the Broad Institute’s MSigDB (Liberzon et al., 2015) and ranked gene lists were tested for enrichment using GSEA’s GSEAPreranked feature, version 3.0. To compare functional annotation enrichment among differentially expressed genes across AD and normal aging, GO terms identified in each comparison (late-stage AD: 14 month AD-BXD vs 14 month Ntg-BXD, **Supplemental Table 7** and normal aging: 6 month Ntg-BXD vs 14 month Ntg-BXD, **Supplemental Table 7**) were extracted from GSEA results. A score for enrichment strength was calculated by transforming the FDR q-values generated by GSEA using the following formula: -log10(FDRq + 0.001), similar to that described in (Raj et al., 2017). The values calculated in each scenario (AD and normal aging) were then plotted against each other in order to identify those pathways with stronger enrichment in AD than normal aging, and vice versa.

**Calculation of a genetic risk score**

To evaluate whether the AD-BXD panel was sensitive to variation in known AD risk loci, we derived a genetic risk score (GRS) for each strain, similar to that described by Chouraki and colleagues in 2016 (Chouraki et al., 2016). Strains were first stratified into impaired (below the AD-BXD population average) and unimpaired (above the AD-BXD population average) based on CFM performance at 6 months of age. We then identified the genotype of each strain at 21 genes known to contribute risk for AD and identified the risk allele of each gene (i.e. the allele that appeared more frequently in the impaired group). Note as some genes appeared in the same linkage block, only 19 genotypes were used in the calculation of the GRS. The odds ratio for each gene was calculated and log transformed to determine an individual risk score per gene. These individual risk scores were used to derive an overall GRS for each strain that reflected how many copies of each risk allele were present. Overall GRS were transformed based on previous methods (Chouraki et al., 2016) using the following formula (**Equation 3-2**):

\[
\text{total risk score} \times (\text{# of markers tested/sum of individual gene risk scores}). (\text{Eq. 3-2})
\]
The GRS was then correlated to cognitive traits as reported. To avoid influence from our original definition of ‘impaired’ versus ‘unimpaired’ using 6 month-old CFM, we correlated GRS to uncorrelated, independent cognitive tasks. As CFC is a cross-sectional task, and we wanted to investigate the extent to which these genes regulated cognitive decline, we focused on cognitive tasks from a separate cohort of aged AD-BXD, particularly 14 month CFA. Finally, to empirically estimate the null distribution for the correlation of our GRS and cognitive traits of interest, we randomly sampled 1000 sets of 19 markers across the genome and repeated the derivation of GRS. We created 1000 GRS from randomly sampled data, and correlated each random GRS to strain-matched cognitive performance, which illustrated the correlation between 14 month CFA and our derived GRS was stronger than the correlation observed for 95% of randomly sampled genes. As an additional control, we repeated the process but based on allelic distribution of risk alleles across Ntg-BXDs. We defined Ntg-BXD strains as ‘impaired’ vs ‘unimpaired’ based on 6 month-old CFM performance, identified the risk allele for each of the 21 genes listed in Table 2, calculated the odds ratio for each, and derived a Ntg-based GRS. This GRS showed no relationship with cognitive outcomes in either Ntg- or AD-BXDs, or any non-cognitive traits tested. Gene lengths were obtained using start and end positions listed in Ensembl version 92 and SNP counts were obtained from Sanger, release REL-1505 (Keane et al., 2011).

Quantification and statistical analyses

All experiments and data analysis were conducted with experimenters blind to strain background and genotype (5XFAD vs Ntg) where appropriate. Statistical analysis was performed using SPSS software Version 23 (IBM), R, and Excel. Distribution was evaluated for normality using Shpiro-Wilk test. Additional analyses included independent unpaired t-tests, univariate ANOVAs, Pearson correlation, and Jaccard index to test similarity. Correction for multiple comparisons was also used where appropriate (i.e differential expression analysis). Data values reported in both the main text and figure legends are given as mean ± standard error of the mean unless otherwise stated. Outliers were identified based on a pre-defined criterion of average values ± 3 SD outside the mean.

Data and software availability

Genotypes from the BXD strains are publicly available on GeneNetwork.org. RNA sequencing from the hippocampus of a subset of AD-BXD strains is available on Gene Expression Omnibus (GEO) under accession number GSE101144. All raw data, including raw phenotype information, have been deposited in the AMP-AD Knowledge Portal synapse at the following link: https://doi.org/10.7303/syn17016211. Strain-averaged behavioral data has been deposited on GeneNetwork.org and is available as Record IDs 20473-20964. EMASE software used for alignment of RNA sequencing reads to a diploid transcriptome is available online at: (https://github.com/churchill-lab/emase)
Results

Genetic background modifies expressivity of FAD mutations

In order to evaluate the influence of genetic background on the impact of causal FAD mutations on behavioral and molecular phenotypes, we generated a panel of 28 genetically diverse F1 mouse strains with and without FAD mutations. Female B6 mice heterozygous for the autosomal dominant 5XFAD transgene (Oakley et al., 2006) were crossed to males from the BXD GRP (Peirce et al., 2004) to generate F1 progeny carrying the 5XFAD transgene (AD-BXDs) or non-transgenic littermates (Ntg-BXDs; Figure 3-1A). Working memory and body weight were monitored bi-monthly and more in-depth phenotyping that included tests of motor function and anxiety was performed at both 6 and 14 months of age (Figure 3-1B). A subset of mice was subsequently tested for long-term spatial learning and memory function using a CFC paradigm (Fanselow, 2000; Neuner et al., 2016). This subset was immediately harvested following CFC testing and tissue was collected for biobanking and later use, including RNA sequencing and enzyme-linked immunosorbent assays (ELISAs) as described below. This time point (immediately following testing) was chosen in order to capture molecular changes corresponding to differences in learning-related intrinsic neuronal excitability reported previously (Kaczorowski and Disterhoft, 2009; Kaczorowski et al., 2011).

As expected (Kaczorowski et al., 2011; Oakley et al., 2006; Ohno, 2009), the 5XFAD transgene accelerated the age at onset (AAO) of working memory deficits in AD-BXD mice relative to Ntg-BXD mice (Figure 3-1C) and exacerbated CFA and CFM deficits, particularly by 14 months of age (Figure 3-1D, E). However, the impact of causal FAD mutations on cognitive performance varied widely depending on the specific background strain evaluated. Notably, this variation in cognitive function parallels the variation observed in human patients harboring FAD mutations (Ryman et al., 2014) and was not correlated with strain-specific variation in activity, pain sensitivity, sensorimotor abilities, or anxiety (Figure B-1). These results suggest the observed variation in cognitive function is regulated, in part, by genetic variants that segregate across the AD-BXD panel. In support, heritability ($h^2_{RB}$) estimates comparing between-strain variance (due to genetic diversity) to total sample variance (due to both genetic and environmental factors), given the average number of biological replicates per strain (Belknap, 1998), demonstrate there is a significant genetic component underlying observed variation (Table 3-1).

Human FAD mutations in APP and PSEN1 included in the 5XFAD transgene increase production of the toxic Aβ42 species, thought to be an initiating factor in a cascade of symptoms eventually leading to neuron loss and dementia (Hardy and Higgins, 1992). To assess the impact of genetic background on the levels of Aβ42 across the panel, brain extracts from 23 AD-BXD strains were assayed in duplicate on human
Figure 3-1. Genetic background modifies AD symptoms in a novel transgenic reference panel.

(A) Female B6 mice heterozygous for the 5XFAD transgene were bred to males from 27 BXD strains to generate genetically diverse but isogenic F1 offspring. (B) Body weight and working memory on the y-maze were measured bi-monthly, and at 6 and 14 months more detailed phenotyping was performed. (C) As expected, onset of working memory deficits was significantly earlier in AD-BXDs compared to Ntg-BXDs [AD-BXDs: n = 223 (123 females/100 males) across 28 strains vs Ntg-BXDs, n = 168 mice (107 female/61 male) across 25 strains, one-tailed t(1,51) = 2.1, p = 0.02]. (D) AD-BXD mice exhibited CFA comparable to Ntg-BXD mice at 6 months [left, one-tailed t(1,48) = 1.4, p = 0.08] but are impaired by 14 months [right, one-tailed t(1,49) = 2.0, p = 0.03] months. Within AD-BXD mouse, background strain significantly modified the impact of the transgene on CFA [effect of strain F(26, 354) = 3.3, p < 0.001]. (E) AD-BXD mice exhibit comparable CFM to Ntg-BXDs at 6 months [left, one-tailed t(1,48) = 1.4, p = 0.08] but are impaired by 14 months [right, one-tailed t(1,49) = 1.9, p = 0.03]. Within AD-BXD mice, background strain significantly modified the impact of the transgene on CFM [effect of strain F(26, 354) = 3.5, p < 0.001]. For D and E, 146 6m AD-BXD (102 females/44 males) and 209 14m AD-BXD (111 females/98 males) across 26 strains were used, along with 114 6m Ntg-BXD (83 females/31 males) across 24 strains and 167 14m Ntg-BXD mice (106 females/61 males) across 27 strains. (F) and (G) Aβ42, as measured by ELISA, increased drastically from 6 to 14 months of age [effect of age F(1,153) = 128.0, p < 0.001] but varied significantly across genetic backgrounds [effect of strain F(22,153) = 2.0, p = 0.01]. n = 154 mice (89 female/65 male) across 23 strains. ND = no data. Strain B6SJL represents the original background strain described in Oakley et al. 2006 for comparison. (H, left) Transgene expression was assessed in subset of AD and Ntg-BXD lines [n = 293 (177 females/116 males across 28 strains)]. Hippocampal RNA sequencing reads were aligned to the human mutant sequence of APP, quantified by TPM and log transformed. AD-BXD mice exhibited significantly greater hAPP expression [t(1,291) = 92.3, p < 0.001]. Across the AD-BXDs, there were no significant effects of age, sex, or (H, right) background strain. Only strains with Aβ42 data are shown here for comparison to (F) and (G). (I, left) Same analysis was done for reads aligned to the mouse endogenous App. Across the panel, 5XFAD mice exhibited slight but significant reduction in App [t(1,291) = 2.6, p = 0.01]. However, within AD-BXD mice there was no effect of age, sex, or (I, right) background strain. For plots C-E, each point represents a strain average. All t-tests in C-E were one-tailed tests based on prior data assessing effects of the 5XFAD transgene on cognitive function (Kaczorowski et al., 2011; Oakley et al., 2006; Ohno, 2009); *p < 0.05.
Table 3-1.  Heritability estimates for phenotypic traits in AD- and Ntg-BXD rats.

<table>
<thead>
<tr>
<th>Trait</th>
<th>Ntg-BXDs</th>
<th>AD-BXD</th>
<th></th>
<th>Heritability ($h^2_{BB}$)</th>
<th></th>
<th>Heritability ($h^2_{BB}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Between-Strain Variance</td>
<td>Av. Within-Strain Variance</td>
<td>Av. n/Strain</td>
<td>Heritability ($h^2_{BB}$)</td>
<td>Between-Strain Variance</td>
<td>Av. Within-Strain Variance</td>
</tr>
<tr>
<td>Age at Onset</td>
<td>7.7</td>
<td>15.7</td>
<td>7.5</td>
<td>0.8</td>
<td>5.7</td>
<td>15.7</td>
</tr>
<tr>
<td>6m CFA</td>
<td>181.2</td>
<td>433.1</td>
<td>5.1</td>
<td>0.7</td>
<td>142.9</td>
<td>293.7</td>
</tr>
<tr>
<td>6m CFM</td>
<td>196.6</td>
<td>340.5</td>
<td>5.1</td>
<td>0.7</td>
<td>163.3</td>
<td>376.8</td>
</tr>
<tr>
<td>14m CFA</td>
<td>146.6</td>
<td>520.7</td>
<td>7.2</td>
<td>0.7</td>
<td>172.0</td>
<td>360.4</td>
</tr>
<tr>
<td>14m CFM</td>
<td>168.0</td>
<td>321.9</td>
<td>7.2</td>
<td>0.8</td>
<td>141.5</td>
<td>299.8</td>
</tr>
<tr>
<td>6m sensorimotor composite</td>
<td>0.3</td>
<td>2.1</td>
<td>8.7</td>
<td>0.6</td>
<td>1.2</td>
<td>3.8</td>
</tr>
<tr>
<td>14m sensorimotor composite</td>
<td>1.7</td>
<td>6.8</td>
<td>7.4</td>
<td>0.6</td>
<td>2.4</td>
<td>9.3</td>
</tr>
<tr>
<td>6m EPM % Time in Open Arms</td>
<td>7.2</td>
<td>62.7</td>
<td>8.7</td>
<td>0.5</td>
<td>38.8</td>
<td>322.7</td>
</tr>
<tr>
<td>14m EPM % Time in Open Arms</td>
<td>29.2</td>
<td>149.6</td>
<td>7.4</td>
<td>0.6</td>
<td>266.6</td>
<td>625.6</td>
</tr>
<tr>
<td>6m amyloid</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>2570.6</td>
<td>2050.5</td>
</tr>
<tr>
<td>14m amyloid</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>36141.9</td>
<td>64897.9</td>
</tr>
</tbody>
</table>

Heritability ($h^2_{BB}$) was determined by calculating the ratio of between-strain variance (i.e. genetic variance) to total sample variance (within-strain variance due to technical/environmental factors plus between-strain variance), given the average number of biological replicates per strain according to established methods (Belknap, 1998). Abbreviations: Av: average, m: months, CFA: contextual fear acquisition, CFM: contextual fear memory, EPM: Elevated Plus Maze, N/A: not applicable.
Aβ42-specific sandwich ELISAs (Oakley et al., 2006). Variation in human Aβ42 levels was heritable (Table 3-1), and overall levels increased with age [effect of age F(1,153) = 128.0, p < 0.001] (Figure 3-1F, G). A significant main effect of strain was observed [F(22,153) = 2.0, p = 0.01], indicating that genetic background significantly modified human Aβ42 levels across the panel. In order to test whether elevated amyloid levels corresponded to an increase in plaque density, we performed IHC analysis on a subset of fixed hemibrains and observed robust plaque deposition in both the hippocampus and cortex of AD-BXD strains, each of which significantly correlated with amyloid levels as measured by ELISA (Figure B-2A, B, and C). As expected, human Aβ42 was not reliably detected in 8 Ntg-BXD brains by ELISA, or in 3 Ntg-BXD brains by IHC (Figure B-2D), suggesting that at least by 6 months of age, Ntg-BXDs do not develop deposition of human Aβ42 compared to their 5XFAD isogenic counterparts. Similar to what is observed in human populations, no significant correlation was observed between amyloid levels and cognitive function (Figure B-3), suggesting partially independent mechanisms work to regulate the extent of cognitive decline and amyloid accumulation.

Differences in cognitive function and Aβ42 pathology were not explained by an effect of age, sex, or background strain on the transcription of the 5XFAD transgene itself, as measured by alignment of RNA sequencing reads from the hippocampus to the mutated human APP (Figure 3-1H, Figure B-4A) or PSEN1 (Figure B-4B) sequences that make up the 5XFAD transgene (Oakley et al., 2006). The lack of a sex difference on either transgene expression or Aβ42 levels is in contrast to a previous report using a single genetic background demonstrating the 5XFAD transgene is differentially expressed based on sex (Sadleir et al., 2015), suggesting that sex-specific effects may vary across genetic backgrounds. In addition, across the AD-BXD panel, there was no effect of genetic background, age, or sex on expression of endogenous App (Figure 3-1I, Figure B-4C) or Psen1 (Figure B-4D). Overall, these results suggest that naturally occurring variants segregating across the AD-B XD panel, rather than artificial differences due to transgene expression, play a significant role in determining susceptibility and/or resilience to changes in cognitive function and amyloid deposition caused by high-risk FAD mutations.

**Cognitive function in the AD-BXDs is sensitive to variation at known AD risk loci**

To test the hypothesis that the inclusion of genetic diversity would better model human AD, we first evaluated whether the AD-BXD panel is sensitive to variation in genes known to confer risk for LOAD. Because the APOE gene is the best characterized risk gene for LOAD in human patients and is relatively well conserved in the mouse (Liao et al., 2015), we queried variants in mouse Apoe. One single nucleotide polymorphism (SNP) in Apoe segregates across the BXD panel (Figure 3-2A), occurring near the receptor-binding region (Mahley et al., 2009). Based on sequence alignment, this SNP causes a switch from glutamate to aspartate at mouse position 163 (Zerbin et al., 2018). While the exact functional consequences of this SNP are unknown, and likely depend on the context of surrounding amino acids, we predicted the D allele of Apoe
Figure 3-2. AD-BXD panel is sensitive to variation in known AD risk gene Apoe. (A) The D allele harbors only a single E > D missense SNP at mouse 163 (red). (B) Across AD-BXD mice, there was a significant effect of Apoe allele [F(1, 354) = 4.7, p = 0.03], age [F(1,354) = 12.3, p = 0.001], and sex [F(1,354) = 17.9, p < 0.001] on CFA. There was a trend toward Apoe having a more significant effect on CFA in females than in males [strain*sex interaction, F(1, 354) = 3.2, p = 0.08]. (C) Across AD-BXD mice, there was a significant effect of Apoe allele [F(1, 354) = 20.9, p < 0.001], age [F(1,354) = 86.2, p < 0.001], sex [F(1,354) = 4.9, p = 0.03], and an age by sex interaction [F(1, 354) = 7.6, p = 0.006] on CFM, demonstrating that while strains carrying the D allele at Apoe do indeed perform more poorly on this task, all female AD mice are more susceptible to AD-related cognitive decline as measured by CFM.
would represent a susceptibility allele across the AD-BXDs based on sequence homology.

To test this hypothesis, we first identified genotyping markers flanking Apoe across the AD-BXDs and then determined the allelic composition of Apoe in each strain. A significant effect of Apoe allele was observed on CFA [F(1,354) = 4.7, p = 0.03], indicating that strains carrying one copy of the D allele of Apoe performed worse on this task (Figure 3-2B). We also observed a significant effect of age [F(1,354) = 12.3, p = 0.001] and sex [F(1,354) = 17.9, p < 0.001] on CFA, as well as a trend toward an interaction between sex and Apoe genotype [F(1,354) = 3.2, p = 0.08]. Together, these results indicate that while most mice exhibited age-related decline in acquisition, female mice generally performed worse on the task and were also particularly susceptible to the effects of the D allele of Apoe. The Apoe effect was even more pronounced when we considered CFM; a significant main effect of Apoe allele was again detected [F(1,355) = 20.9, p < 0.001], along with significant effects of sex [F(1,355) = 4.9, p = 0.03], age [F(1,355) = 86.2, p < 0.001], and a sex by age interaction [F(1,355) = 7.6, p = 0.006] (Figure 3-2C). These results indicate first that mice harboring a single copy of the D allele of Apoe exhibited poorer CFM, and second that female mice are more susceptible to AD-related cognitive decline with age. No effect of Apoe genotype was observed on working memory traits. Across Ntg-BXDs, Apoe genotype exhibited either a less robust effect, or no effect, on cognitive performance on CFA and CFM tasks, respectively (Figure B-5).

Overall, the above data demonstrates that variation at the Apoe locus in mice, particularly those harboring the 5XFAD transgene, is associated with cognitive outcomes. In humans, additional genes have been identified that play small, although significant, roles in regulating susceptibility to AD (Lambert et al., 2013). Recent studies suggest that information about genetic variation at these additional loci, in the form of a GRS, can better predict an individual’s risk of developing AD (Chouraki et al., 2016). In order to evaluate whether naturally occurring variants in genes associated with LOAD risk in humans are associated with cognitive outcomes in the AD-BXD panel, we computed a GRS for each of our strains similar to the method described by Chouraki and colleagues in 2016 (Chouraki et al., 2016). First, we stratified strains into impaired and unimpaired groups based on 6 month-old CFM (Figure 3-3A). We then identified the genotype of each strain at 21 LOAD risk genes (across 19 genetic markers, Table 3-2) and classified the risk allele of each gene as that allele which appeared more frequently in the impaired group. Odds ratios were calculated and transformed based on risk allele dosage to obtain a final GRS for each strain, which was normally distributed across the panel (Shapiro-Wilk test for normality p = 0.7, Figure 3-3B).

Once each GRS was calculated, we then asked how well a strain’s score predicted cognitive outcomes as measured on an uncorrelated task in a separate cohort of AD-BXD mice (i.e. 14 month-old CFA). Although no individual risk gene significantly differentiated impaired vs unimpaired strains at 6 months, when taken together, the GRS was significantly associated with cognitive outcomes in AD-BXD mice (Figure 3-3C).
Figure 3-3. Genetic risk score calculated from genotype at known AD risk genes predicts cognitive decline.

(A) Strains were stratified into impaired (below population average) and unimpaired (above population average) based on 6m CFM performance. (B) GRS were calculated for each strain based on allelic composition of 21 genes known to confer risk for AD. The risk allele of each gene was defined as that which appeared more frequently in the impaired population pictured in (A). (C) GRS significantly predicts how a given AD-BXD strain will perform on CFA at 14m. (D) No relationship between GRS and CFA in 14m Ntg-BXD mice was observed. (E) No relationship between GRS and non-cognitive traits across AD-BXDs including 14m Aβ42 load, (F) 14m weight, (G) 14m sensorimotor performance, (H) or 14m anxiety as measured by percent open entries on the EPM was observed.
Table 3-2. Genes known to confer risk for human AD and their odds ratio across the AD-BXDs.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Mouse Chr.</th>
<th>SNP Density (SNP/Kb)</th>
<th>High-Impact Changes</th>
<th>Risk Allele</th>
<th>Odds Ratio</th>
<th>95 % CI</th>
<th>Z-stat</th>
<th>Pval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inpp5d</td>
<td>1</td>
<td>1.24</td>
<td>NMD, SNP, indel</td>
<td>B</td>
<td>1.30</td>
<td>0.28 - 6.3</td>
<td>0.36</td>
<td>0.72</td>
</tr>
<tr>
<td>Cr1l</td>
<td>1</td>
<td>0.06</td>
<td></td>
<td>D</td>
<td>2.00</td>
<td>0.41 - 9.8</td>
<td>0.85</td>
<td>0.39</td>
</tr>
<tr>
<td>Celf1</td>
<td>2</td>
<td>0.70</td>
<td></td>
<td>D</td>
<td>1.50</td>
<td>0.30 - 7.4</td>
<td>0.50</td>
<td>0.62</td>
</tr>
<tr>
<td>Cass4</td>
<td>2</td>
<td>0.08</td>
<td></td>
<td>D</td>
<td>1.50</td>
<td>0.30 - 7.4</td>
<td>0.50</td>
<td>0.62</td>
</tr>
<tr>
<td>Zcwpwl</td>
<td>5</td>
<td>0.11</td>
<td></td>
<td>D</td>
<td>1.63</td>
<td>0.34 - 8.0</td>
<td>0.61</td>
<td>0.54</td>
</tr>
<tr>
<td>Epha1</td>
<td>6</td>
<td>0.00 (indel)</td>
<td></td>
<td>D</td>
<td>1.60</td>
<td>0.33 - 7.8</td>
<td>0.58</td>
<td>0.56</td>
</tr>
<tr>
<td>Cd33</td>
<td>7</td>
<td>3.17</td>
<td>MS, Stop gained</td>
<td>D</td>
<td>1.67</td>
<td>0.30 - 9.2</td>
<td>0.59</td>
<td>0.56</td>
</tr>
<tr>
<td>Picalm</td>
<td>7</td>
<td>1.75</td>
<td></td>
<td>D</td>
<td>3.60</td>
<td>0.71 - 18.3</td>
<td>1.55</td>
<td>0.12</td>
</tr>
<tr>
<td>Sorl1</td>
<td>9</td>
<td>12.49</td>
<td>MS, SRV, SAV</td>
<td>D</td>
<td>2.50</td>
<td>0.50 - 12.6</td>
<td>1.11</td>
<td>0.27</td>
</tr>
<tr>
<td>Abca7</td>
<td>10</td>
<td>0.05</td>
<td></td>
<td>D</td>
<td>1.17</td>
<td>0.24 - 5.6</td>
<td>0.19</td>
<td>0.85</td>
</tr>
<tr>
<td>Slc24a4</td>
<td>12</td>
<td>4.07</td>
<td>SRV</td>
<td>D</td>
<td>3.60</td>
<td>0.71 - 18.3</td>
<td>1.55</td>
<td>0.12</td>
</tr>
<tr>
<td>Rin3</td>
<td>12</td>
<td>4.95</td>
<td>MS</td>
<td>B</td>
<td>0.86</td>
<td>0.18 - 4.1</td>
<td>0.19</td>
<td>0.85</td>
</tr>
<tr>
<td>Mef2c</td>
<td>13</td>
<td>0.15</td>
<td></td>
<td>B</td>
<td>0.86</td>
<td>0.18 - 4.1</td>
<td>0.19</td>
<td>0.85</td>
</tr>
<tr>
<td>Nme8</td>
<td>13</td>
<td>3.68</td>
<td>SRV</td>
<td>D</td>
<td>1.40</td>
<td>0.30 - 6.6</td>
<td>0.42</td>
<td>0.67</td>
</tr>
<tr>
<td>Clu</td>
<td>14</td>
<td>0.00 (indel)</td>
<td>NMD</td>
<td>D</td>
<td>5.50</td>
<td>0.84 - 36.2</td>
<td>1.77</td>
<td>0.08</td>
</tr>
<tr>
<td>Ptk2b</td>
<td>14</td>
<td>1.85</td>
<td>MS, SRV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ferm2</td>
<td>14</td>
<td>2.40</td>
<td></td>
<td>D</td>
<td>1.83</td>
<td>0.32 - 10.6</td>
<td>0.68</td>
<td>0.50</td>
</tr>
<tr>
<td>Cd2ap</td>
<td>17</td>
<td>4.85</td>
<td>MS, SRV</td>
<td>-</td>
<td>1.00</td>
<td>0.21 - 4.7</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>H2-Eb1</td>
<td>17</td>
<td>21.41</td>
<td>MS, SRV, stop gained</td>
<td>D</td>
<td>1.05</td>
<td>0.22 - 5.0</td>
<td>0.06</td>
<td>0.95</td>
</tr>
<tr>
<td>Trem2</td>
<td>17</td>
<td>0.13</td>
<td></td>
<td>B</td>
<td>1.20</td>
<td>0.25 - 5.8</td>
<td>0.23</td>
<td>0.82</td>
</tr>
<tr>
<td>Bin1</td>
<td>18</td>
<td>0.21</td>
<td>MS</td>
<td>D</td>
<td>1.33</td>
<td>0.28 - 6.3</td>
<td>0.36</td>
<td>0.72</td>
</tr>
</tbody>
</table>

Abbreviations: SNP: single nucleotide polymorphism, indel: insertion/deletion, Chr: chromosome, SRV: splice region variant, SAV: splice acceptor variant, MS: missense
Notably, the GRS was not associated with cognition in Ntg-BXDs, suggesting genes used to create the GRS exhibit more specificity toward mediating AD-related decline (Figure 3-3D). We repeated this entire process with 1,000 sets of 19 randomly selected genetic markers and determined the correlation of the GRS and 14 month AD-BXD CFA was among the top 5% of all observed permutations, suggesting the additive association of LOAD risk genes with 5XFAD-related cognitive decline is much greater than a set of genes randomly distributed across the genome. In addition, a GRS derived from genotypes at the same risk alleles, but using the distribution of ‘impaired’ and ‘unimpaired’ Ntg-BXD strains, rather than AD-BXD strains, to define odds ratios for each individual LOAD risk gene showed no relationship with late-disease cognitive outcomes in either 14 month-old Ntg-BXDs or AD-BXDs (Figure 2-6), further demonstrating these genes uniquely interact with the 5XFAD transgene. Finally, the original GRS (Figure 3-3B) showed no association to non-cognitive traits such as Aβ42 levels, weight, sensorimotor abilities, or anxiety (Figure 3-3E-H). Overall, these results demonstrate 1) the AD-BXD panel is sensitive to variation in known LOAD risk variants, and 2) the CFA task is particularly sensitive to this variation, and has the potential to be used as a translationally relevant cognitive assay in preclinical AD studies.

**AD-BXD transcriptome shows concordance with late-onset AD signature**

We next decided to investigate whether or not the AD-BXD panel shared similarities with human AD at the transcriptional level. We first performed RNA sequencing on hippocampal tissue from a subset of AD-BXDs and Ntg-BXDs and evaluated the expression of genes known to be misregulated in AD. As expected from studies of post-mortem human tissue, the 5XFAD transgene significantly altered the expression of a number of these genes, particularly *Bin1*, *Clu*, *Cd33* (Karch et al., 2012), *Trem2* (Piccio et al., 2016), and *C1qa* (Hong et al., 2016) (Figure 3-4A). Similar to what we observed for behavioral and pathological phenotypes, risk gene expression varied across the AD-BXD panel. This suggests genetic background may influence AD susceptibility by altering underlying transcriptional networks, so to gain a mechanistic understanding of functional categories altered in AD-BXDs relative to Ntg-BXDs, we performed differential expression analysis using DESeq2 (Love et al., 2014) followed by GSEA (Subramanian et al., 2005, Supplemental Table 6). As expected, the GO functional categories most significantly enriched among genes observed to be down-regulated in AD largely related to neuronal activity, structure, and function (Figure 3-4B, left) while the GO functional categories most significantly enriched among genes observed to be up-regulated in AD related largely to immune response (Figure 3-4B, right). Together, these data highlight the maintenance of neuron activity, particularly the activity of select ion channels and receptors, as pathways that may be augmented to promote resilience, while immune pathways as those that may need to be suppressed to promote resilience.

To further evaluate whether observed changes in our AD-BXD model paralleled those observed in human patients, we next performed a series of cross-species
Figure 3-4. Genetic background modifies AD-associated transcriptome. (A) Genes known to be associated with AD are differentially expressed in our panel, n = 132 mice (65 females/67 males across 15 strains). Each point represents a single genotype/strain/age/sex averaged sample, **p < 0.05 two-tailed t-test. (B) List of top GO functional categories enriched among genes (left) down-regulated and (right) up-regulated in all AD-BXDs relative to all Ntg-BXDs.
comparative analyses using aged brain tissue (14 month AD-BXD mice) to best parallel the tissue available from human patients. First, we evaluated the expression of a set of 60 core genes previously defined as a human AD consensus signature, primarily enriched for down-regulated mitochondrial and neuronal genes (Supplemental Table 5, Hargis and Blalock, 2017). We observed higher concordance between our mouse panel and this human AD signature (Figure 3-5A) than that reported for other AD models on a single genetic background (Hargis and Blalock, 2017). This effect replicated in 3 independent human datasets tested (Figure B-7). Second, we noted that the significant upregulation of immune-related pathways in our AD-BXD mice (Figure 3-4B) paralleled the significant association of immune-related genes with human AD, both at the transcriptional and genetic level (International Genomics of Alzheimer's Disease, 2015; Zhang et al., 2013). To test whether the identity of genes driving this association were similar across mice and humans, we used GeneWeaver (Baker et al., 2016) to calculate overlap of genes up-regulated in aged 14 month-old AD-BXD mice (Supplemental Table 7) and two gene lists associated with human AD. First, we utilized a list of genes belonging to the transcriptional co-expression module most highly associated with human AD identified by Zhang and colleagues (Zhang et al., 2013), and second, a list of 151 highly connected AD-related genes identified by Jones et al. (International Genomics of Alzheimer's Disease, 2015). Each of these lists were significantly enriched for genes with immune-related annotations. In both cases, the overlap between mouse and human signatures was significant (Figure 3-5B).

Finally, we tested whether the AD-specific enrichment of immune-related pathways observed in human AD, but not normal aging (Raj et al., 2017), was preserved across our AD and Ntg-BXDs. To do this, we identified GO terms enriched among those genes significantly differentially expressed between 14 month-old AD and Ntg-BXDs (Supplemental Table 7, 5XFAD-related genes) and those enriched among genes significantly differentially expressed between 6 month and 14 month Ntg-BXDs (Supplemental Table 7, normal aging-related genes). To enable comparison across datasets, we identified those GO terms with enough genes to be identified in each set and compared enrichment strength across AD and normal aging in our mouse panel (Figure 3-5C and Supplemental Table 8). Enrichment of immune-related terms was exclusively observed among our list of 5XFAD-related genes, and not normal aging-related genes. A similar trend was observed in neuron and ion-channel related terms, suggesting downregulation of neuron structure, function, and/or activity to also be a unique feature of AD relative to normal aging in the mouse. Changes unique to normal aging include DNA metabolism, RNA processing, and peptidase activity (Figure 3-5C, bottom right). Overall, the incorporation of genetic diversity into a mouse model of AD resulted in a transcriptomic profile that more closely matched human AD than previous AD models with limited genetic background variation (Hargis and Blalock, 2017).
Figure 3-5. Aged AD-BXD transcriptome shows high concordance with late-onset human AD signature.
(A) 39 out of 60 (65%) of AD signature genes identified by and Blalock (Hargis and Blalock, 2017) show concordant significant changes in expression across mouse and human transcriptomes. The log₂ fold change (FC) of significantly differentially expressed genes between 14 month-old AD- and Ntg-BXDs is plotted on the y-axis, while the log₂FC of gene expression between human AD patients and controls from study by a study by Berchtold and colleagues (Berchtold et al., 2013) is plotted on the x-axis. Each point represents a single gene; discordant genes with log₂ fold changes with opposite direction have been highlighted in red. (B) Genes up-regulated in 14 month AD-BXDs relative to Ntg-BXDs were compared to genes associated with human AD by (top) Zhang and colleagues (Zhang et al., 2013) and (bottom) Jones and colleagues (International Genomics of Alzheimer's Disease, 2015). A significant overlap was identified in both cases. (C) Graph of enrichment strength of GO categories across (y-axis) mouse AD or (x-axis) normal aging. GSEA was performed on genes identified to be differentially expressed relative to 5XFAD carrier status or normal aging in Ntg-BXDs (Supplemental Table 7). For GO terms that were identified in both scenarios, the FDR q-values were transformed to obtain a measure of enrichment strength and scores were plotted against each other to identify unique and/or common differentially regulated GO terms. As such, each axis can be thought of significance; the upper left quadrant highlights pathways that are uniquely significantly altered in AD-BXDs relative to Ntg-BXDs, while the bottom right quadrant highlights pathways that are uniquely significantly altered in normal aging (14 month Ntg-BXDs vs 6 month Ntg-BXDs). Data points are colored based on directionality of enrichment score calculated by GSEA: red = genes belonging to this category were significantly up-regulated in given scenario, blue = genes belonging to this category were significantly down-regulated in given scenario. Dotted lines represent enrichment scores for FDR q-value = 0.05.
Discussion

**AD-BXD panel represents a new translational model of human AD**

It has long been recognized that Alzheimer’s disease is a complex and polygenic disease, likely influenced by multiple variants, some with relatively small effect sizes (Lambert et al., 2013). Here, we introduce the first genetically diverse population of AD mice - the AD-BXDs - as a more translational model of human AD by demonstrating a high level of concordance between the AD-BXDs and both familial and sporadic forms of human AD at the molecular and behavioral level. In particular, the observed variation in AAO mirrors the variation in human patients reported by Ryman and colleagues (Ryman et al., 2014), suggesting our panel captures a portion of the phenotypic heterogeneity observed in human FAD patients. In addition, our female AD-BXD mice appear to be more susceptible to Apoe risk and AD-related cognitive decline (Figure 3-2B-C) despite comparable levels of Aβ42 deposition (Figure 3-1F-G), 5XFAD transgene expression (Figure 3-1H), and endogenous App levels (Figure 3-1I), similar to epidemiological trends observed in human patients (Altmann et al., 2014; Mielke et al., 2014; Zokaei et al., 2017). At the genetic level, we demonstrate that the extent of AD-related cognitive decline is influenced by a given strain’s specific allele distribution across a set of 21 loci associated with sporadic LOAD. Transcriptionally, the changes occurring in the AD-BXDs relative to Ntg-BXDs, particularly at aged time points, show a high level of overlap with transcriptional changes occurring in human AD patients relative to age-matched controls, both in terms of up-regulated inflammatory pathways and down-regulated neuronal signatures, suggesting some common molecular mechanisms exist between mouse and human. Overall, results here demonstrate the critical role genetic background plays in determining susceptibility to disease and present the AD-BXD panel as a useful tool that will enable the identification of modifier genes more likely to translate to human patients.

**Genetically diverse isogenic mice as a resource for experimental precision medicine**

In recent years, there has been growing skepticism regarding the utility of AD mouse models, in part because research using these models have failed to translate into successful treatments. There are a variety of reasons that may explain this failure, many of which have been discussed previously (Onos et al., 2016). However, a common theme of many traditional models is a lack of genetic diversity. Regardless of the specific strain used, the use of a single genetic background precludes the ability to understand the impact of individual genetic diversity on disease trajectory. Here we introduce the AD-BXD panel as a new preclinical resource that will allow future studies to investigate the influence of genetic complexity on the behavioral, molecular, and pathological phenotypes in AD, as well as response to interventions. By layering multiple scales of data collected from the AD-BXDs (phenotype, transcriptomics, proteomics, etc.), we may ultimately be able to identify subtypes of AD and begin to develop personalized therapeutic interventions. Given the advantages represented by this experimental design,
and the increasing awareness that genetic background impacts a variety of complex traits (Sittig et al., 2016), the approach used here should be of broad interest across scientific disciplines. We hypothesize the incorporation of genetic diversity into preclinical studies of various complex diseases will greatly enhance the overall translational potential of mouse models. In addition, our experimental design is likely to be broadly applicable to mouse models of human disease that incorporate a dominantly inherited high-risk genotype in the form of a transgene or other genetic perturbation. In these cases, the ultimate identification of genetic factors that modify disease onset and/or severity will provide insight into pathways critical for regulating disease pathogenesis.

**The role of modifier genes in normal aging and resilience to AD**

The creation of the genetically diverse Ntg and AD-BXD panel enables the use of genetic mapping to identify modifier alleles that influence the onset and severity of cognitive decline in both normal aging and AD. The inclusion of both Ntg and AD lines is particularly powerful, as the extent to which genetic mechanisms that underlie normal cognitive aging and AD overlap is still unclear. Comparison of mapping results across genotypes will identify alleles that either act as general modifiers and contribute to a phenotype regardless of disease status, or specific AD modifiers that exhibit an epistatic relationship with the 5XFAD transgene. To this end, our results indicate that the B6 background strain may contribute modifiers that increase resilience to high-risk 5XFAD mutations, creating an ideal opportunity to model resilience for the first time. In particular, the B6 background appears to attenuate the impact of 5XFAD mutations on cognitive traits, despite moderate-to-high levels of Aβ42 (Figure 3-1G). In addition, across 21 genes known to confer risk for LOAD, the B allele represents the protective allele in 17 cases (Table 3-2). Future studies will utilize these resources to perform genetic mapping to identify genomic regions involved in the regulation of cognitive decline, and precise genes present in the B6 background that contribute to resilience. Due to large haplotype blocks segregating among the BXD strains, it is likely additional Ntg- and AD-BXD strains will need to be incorporated to narrow in on causal variants, genes, and pathways regulating quantitative traits across the panel. As the functional interrogation of all candidate genes will require large investments of time and resources, it is likely the development of new disease-modifying strategies will be a community effort.

**Improving rigor and reproducibility in preclinical studies**

In addition to providing an ideal preclinical resource to identify modifiers of susceptibility and resilience to AD, the AD-BXD panel presents a multitude of additional advantages. Notably, as each parental line (B6 and BXDs) is fully inbred, the F1 mice described here can be recreated across time and laboratories, maximizing the utility of our characterization of individual lines as either cognitively resilient or susceptible and enhancing rigor and reproducibility of the approach. The use of genetically identical, or isogenic, F1 mice also allows for calculation of heritability across a number of diverse
traits, ranging from body weight and sensorimotor abilities to anxiety and cognitive function. Traits with high heritability are, by definition, largely influenced by genetics rather than environmental or technical factors and thus are more likely to replicate across time and laboratories. Identification of traits that are robust to environmental influences will enhance reproducibility of research and confidence that results will translate from the bench to clinic more rapidly. Traditional studies that utilize only a single inbred strain to study disease are useful for understanding basic mechanisms associated with symptom onset, but are equivalent to study disease in a single human. In contrast, results from studies that utilize diverse genetic backgrounds are a better model of complex disease across individuals, and are therefore, more likely to generalize across patient populations.
CHAPTER 4. IDENTIFICATION OF GENES AND NETWORKS THAT MODIFY SUSCEPTIBILITY TO ALZHEIMER’S DISEASE*

Introduction

AD is the most common neurodegenerative disease, characterized by a combination of severe memory impairment and two classical neuropathologies, extracellular amyloid plaques and intracellular neurofibrillary tangles (Selkoe, 1991). According to the amyloid hypothesis, which has been extensively researched for decades, the formation and deposition of amyloid, particularly toxic Aβ42, is thought to be an initiating factor that leads to later neurodegeneration and cognitive impairment (Hardy and Selkoe, 2002; Selkoe and Hardy, 2016). However, many imaging and post-mortem studies of human brains have shown that substantial amounts of AD pathology, particularly plaque pathology, can be present in the brains of cognitively intact individuals (Morris et al., 1996; Negash et al., 2013). These individuals, who often meet the criteria for a pathological diagnosis of AD but remain asymptomatic, represent a clinically interesting subset of the population that exhibit a certain degree of resilience to what are typically highly deleterious neuropathologies.

Resiliency, defined as better cognitive functioning than predicted based on given pathological, genetic, or molecular characteristics, has been observed in both the general population (Hohman et al., 2016) and in families harboring high-risk genetic mutations that often confer early-onset or FAD (Ryman et al., 2014). While FAD is typically thought to be a severe form of the disease with an age of onset before 65, there is a wide range in the age at first symptom onset (Ryman et al., 2014). In both sporadic LOAD and FAD, disease onset is highly heritable (Gatz et al., 1997), indicating genetic factors likely play a large role in determining individual susceptibility or resilience. Identifying these specific genetic factors, particularly those that promote resilience, would provide key insight into disease mechanisms and nominate putative targets for therapeutic intervention, as strategies that delay disease onset even by a few years would provide much needed disease modifying therapies.

Despite the immense therapeutic potential presented by resilience factors, their identification in human populations remains elusive. Several studies have utilized FAD populations to identify modifiers of the age at first symptom onset (Lee et al., 2015), but these populations are typically not large enough to support genome-wide testing and identification of resilience factors. In addition, the identification of individuals with little to no family history of AD and who remain cognitively intact despite high pathology

loads is almost impossible in the general population, as these individuals rarely enter the clinic. Even when resilient individuals enroll as part of an observational study, human-oriented research presents additional challenges. The human genome is incredibly complex, individuals are consistently exposed to a variety of uncontrollable environmental factors, and accessing critical disease-relevant tissue at early stages of disease is uncommon.

Here we take advantage of the inbred nature of the AD-BXD panel described in Chapter 3 to identify 1) genes associated with cognitive decline and neuropathological symptoms across the panel, and 2) transcriptional networks present at early stages of disease (6 months) that predict cognitive impairment later in disease (14 months). Genetic mapping identified a number of genome-wide significant quantitative trait loci (QTLs) containing variants that directly modify susceptibility to AD-related cognitive and pathological traits, providing a valuable list of candidate genes that may be targeted as therapeutics to prevent or delay the onset of disease. In addition, at 6 months of age, the AD-BXDs are cognitively unimpaired as a population relative to their non-transgenic littermates (Neuner et al., 2019a) as measured by CFC, making the 6-month time point ideal to profile networks present prior to overt symptom onset. Understanding the molecular mechanisms that occur early in disease may help to identify causal drivers of disease pathogenesis and therapeutic targets for interventions. As there is currently no cure for AD, work here is poised to contribute significantly to human health.

Methods

Animals

All data used in this study came from mice that were part of the AD-BXD panel, which has been previously described (Neuner et al., 2019a). As both of these resources (B6-5XFAD and individual BXD strains) consist of fully inbred mice, one generation of breeding results in isogenic F1 AD-BXD mice that harbor the 5XFAD transgene in combination with a genetically diverse BXD chromosome. As female 5XFAD mice are hemizygous, non-transgenic F1 mice were also generated (in approximately 50/50 ratio). Male and female AD-BXD mice were group housed with a mix of transgenic and non-transgenic same-sex littermates and maintained on a 12 hour light/dark cycle. All mouse experiments occurred at University of Tennessee Health Science Center and were carried out in accordance with the standards of the Association for the Assessment and Accreditation of Laboratory Animal Care (AAALAC), as well as the recommendations of the National Institutes of Health Guide for the Care and Use of Laboratory Animals. The protocol was approved by the Institutional Animal Care and Use Committee (IACUC) at the University of Tennessee Health Science Center.
Genetic mapping

Mapping of behavioral and pathological traits

For males and females of each AD-BXD strain, averages were calculated for 1) age at onset, 2) CFA at 6 and 14 months, 3) CFM at 6 and 14 months, and 4) Aβ42 levels at 6 and 14 months to use for genetic mapping using the Haley-Knott regression analysis in r/qtl (Broman et al., 2003). Genotypes for BXD strains (mm10) were downloaded from GeneNetwork.org. R/qtl’s scanone function using Haley-Knott regression and a single QTL model was used to identify regions of the genome containing variants responsible for regulating observed variation in quantitative traits (Broman et al., 2003). Sex and age were used as covariates where appropriate to allow for the possibility of different genetic contributions between sexes. Sex was included as a covariate if it was observed to have a main effect on the trait by ANOVA, or interact with our variables of interest (i.e. age and genotype). Permutation tests were used to evaluate significance. The recommended 1.5 LOD confidence interval was identified and lists of genes located in the interval were obtained using the R package biomaRt v2.34.2 (Durinck et al., 2009).

Mapping of transcriptional traits

Genotypes for BXD strains were obtained from GeneNetwork.org (Mulligan et al., 2017). Module eigengenes as generated by weighted gene co-expression network analysis (WGCNA) were exported and used as quantitative traits for downstream QTL mapping in r/qtl (Broman et al., 2003). Sex was used as an additive covariate and permutation tests were used to determine statistical significance.

Transcriptome analyses

RNA sequencing

Initial RNA sequencing from hippocampus of the AD-BXD panel [5XFAD positive only; 6 months, n = 33 (15 females/18 males) and 14 months n = 36 (16 female/20 male)] has been previously reported (Neuner et al., 2019a), GEO accession number GSE101144]. Here, we expand upon this dataset and include RNA sequencing data from an additional 38 (32 female/6 male) 6 month and 50 (29 female/21 male) AD-BXD mice, for a total of 157 AD mice, 71 of which were 6 months of age (47 female/24 male) and 86 of which were 14 months of age (45 female/41 male). RNA sequencing was performed as previously described (Neuner et al., 2019a). Briefly, all samples were isolated using the Qiagen RNeasy Mini kit, libraries were prepared using Truseq Stranded mRNA Sample Preparation Kit (Illumina Inc), and sequenced by 75bp sequencing on an Illumina HiSeq2500. The GBRS pipeline (https://github.com/churchill-lab/gbrrs; Choi et al., 2017) was used to first align reads to a diploid B6/D2 transcriptome using Bowtie followed by an expectation maximization algorithm (Raghupathy, et al. 2018) to quantify the number of reads aligned to either the B6 or D2 allele. This allows for the quantification of both total reads assigned to a gene and the number of reads
assigned to either the $B$ or $D$ allele. Genes were filtered to require an average of at least 1 transcript per million (TPM) in 50% of samples, RNA data was batch corrected using ComBat (Johnson et al., 2007; Leek et al., 2012), and biological replicates were averaged together for downstream analyses. Specifically, samples from individual mice from the same strain, sex, and age were averaged together to derive one group average. Data from both GSE101144 and new data reported here (available now on GEO as GSE119215) represents data from 79 strain/sex/age groups across 28 background strains. Differential expression analysis comparing strain/sex averaged 6 month AD-BXD and 14 month AD-BXD gene expression was performed using DESeq2 (Love et al., 2014).

**Gene set enrichment analysis**

GSEA was performed according to established procedures (Liberzon et al., 2011; Subramanian et al., 2005). Genes significantly differentially expressed relative to age in the AD-BXDs (14 month versus 6 month, adjusted p-value ≤ 0.05) were sorted by log2 fold change and uploaded to GSEA desktop software. Using the “GSEAPreranked” tool and the MSigDb 3.0, GO terms from all categories (Biological Process, Molecular Function, and Cellular Compartment) significantly enriched among either down-regulated or up-regulated genes were identified.

**Cell-type specific enrichment analysis**

For cell-type specific enrichment analysis, a list of genes and their cell-type assignment based on max FPKM from single-cell RNA sequencing in the mouse cerebral cortex was obtained from brainrnaseq.org (Zhang et al., 2014). A hypergeometric test was used to determine the statistical significance of overlap between our list of differentially expressed genes and lists of cell-specific genes. Only genes from differential expression that had a cell-type assignment were used; as such, the list of 12,978 genes downloaded from brainrnaseq.org was used as our background gene set for calculating statistical significance of overlap.

**Weighted gene co-expression network analysis**

WGCNA was performed according to established methods (Langfelder and Horvath, 2008). Additional filtering required at least 1 TPM in 20% of samples, for a final gene list containing ~16,000 genes. A minimum module size of 30 was implemented, and block-wise network construction was used to assemble modules using only 6 month-old AD-BXD RNA data. The WGCNA function GOenrichmentAnalysis was used to identify GO terms significantly enriched within each of the modules using a false discovery rate of 0.05. Hub gene identification using partial correlation analysis was performed using the Statistical Inference of Large-Scale Gaussian Graphical Model in Gene Networks (SILGGM) package in R (Zhang et al., 2018). Module eigengenes representing the first principal component of the expression matrix of the corresponding module were derived using standard methods within the WGCNA package (Langfelder and Horvath, 2008). These eigengenes were used as representative measures of gene expression profiles within a given module and are represented using arbitrary
standardized units. To evaluate how increased module eigengene expression related to actual module member expression, we also derived a standardized module expression value from the mean expression level of all genes as previously reported (Mostafavi et al., 2018).

**Data and software availability and statistical analysis.**

Genotypes from the BXD strains are publicly available on GeneNetwork.org. All raw data, including raw phenotype information, have been deposited in the AMP-AD Knowledge Portal synapse at the following link: https://doi.org/10.7303/syn17016211. New RNA sequencing data reported here is available on GEO as GSE119215. Strain-averaged behavioral data has been deposited on GeneNetwork.org and is available as Record IDs 20473-20964. R software version 3.4.3 was used for data analysis. WGCNA version 1.63 was used for network analysis. Statistical tests included paired t-tests, ANOVA, hypergeometric tests for overlap, Spearman’s correlation, and permutation tests. Data are reported here as mean ± standard error unless otherwise stated.

**Results**

**Identification of novel loci modifying AD susceptibility**

Having established that the AD-BXD panel is a good model for human AD at the phenotypic, genetic, and transcriptional level, and that individual genetic variation plays a significant role in mediating susceptibility or resilience to high-risk AD mutations, we next sought to identify precise genomic regions modifying cognitive and pathological traits. In order to do this, we performed genome-wide QTL mapping using r/qtl. Strain or strain/sex averaged data was used where appropriate. Significant genome-wide QTLs on chromosome 3 and chromosome 2 were identified for CFA and CFM in the AD-BXDs, respectively (Table 4-1), each containing a number of positional candidates (Supplemental Table 9). In addition, a significant genome-wide QTL on a distinct region of chromosome 3 was identified as harboring specific modifiers of Aβ42 levels at 14 months (Table 4-1), which also contained a number of positional candidate genes (Supplemental Table 9). Notably, this region on chromosome 3 did not overlap with identified cognitive QTLs, suggesting independent genetic mechanisms mediate susceptibility or resilience to the cognitive and pathological symptoms of AD. In addition, this QTL was only significant when considering a sex-interactive model, and genotype at the peak QTL marker was determined to significantly interact with sex to influence Aβ42 levels (p < 0.05), whereby the peak QTL had a more pronounced effect in female AD-BXD mice than male AD-BXD mice. Although no significant QTLs were identified as mediating the AAO for AD-BXD mice, given its high heritability, it is likely that AAO is a complex trait mediated by multiple small-effect variants unable to be identified with our current sample size.
Table 4-1. List of QTLs reaching genome-wide significance across the AD-BXD and Ntg-BXD populations.

<table>
<thead>
<tr>
<th>Trait</th>
<th>Input</th>
<th>Covariates</th>
<th>Model</th>
<th>Peak LOD</th>
<th>Peak Location</th>
<th>Peak Chr</th>
<th>1.5 LOD Start (Mb)</th>
<th>1.5 LOD End (Mb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CFA</td>
<td>All Data</td>
<td>Age, Sex, &amp; Genotype</td>
<td>Additive</td>
<td>4.58</td>
<td>rs31878001</td>
<td>5</td>
<td>18.78</td>
<td>144.40</td>
</tr>
<tr>
<td>CFM</td>
<td>5XFAD</td>
<td>Sex &amp; Age</td>
<td>Additive</td>
<td>3.73</td>
<td>rs13476358</td>
<td>2</td>
<td>13.44</td>
<td>162.69</td>
</tr>
<tr>
<td></td>
<td>All Data</td>
<td>Age, Sex, &amp; Genotype</td>
<td>Additive</td>
<td>8.04</td>
<td>rs13476363, rs27125935, rs6384589</td>
<td>2</td>
<td>16.31</td>
<td>42.56</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Interactive</td>
<td>8.23</td>
<td>rs27125935, rs6384589, rs13476363</td>
<td>2</td>
<td>16.31</td>
<td>44.40</td>
</tr>
<tr>
<td></td>
<td>Ntg</td>
<td>Sex &amp; Age</td>
<td>Additive</td>
<td>4.9</td>
<td>c6.loc30</td>
<td>6</td>
<td>51.31</td>
<td>63.79</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4.39</td>
<td>rs13476363</td>
<td>2</td>
<td>14.40</td>
<td>42.56</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4.29</td>
<td>rs33497497</td>
<td>5</td>
<td>19.14</td>
<td>29.54</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Interactive</td>
<td>3.64</td>
<td>rs51257993</td>
<td>14</td>
<td>61.21</td>
<td>99.74</td>
</tr>
<tr>
<td></td>
<td>5XFAD</td>
<td>Sex &amp; Age</td>
<td>Interactive</td>
<td>6.13</td>
<td>c6.loc30</td>
<td>6</td>
<td>50.68</td>
<td>53.96</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6.09</td>
<td>rs3699934</td>
<td>2</td>
<td>35.40</td>
<td>44.48</td>
</tr>
<tr>
<td></td>
<td>Amyloid</td>
<td>14m</td>
<td>Interactive</td>
<td>6.17</td>
<td>rs3699934</td>
<td>2</td>
<td>35.40</td>
<td>44.48</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sex</td>
<td>Interactive</td>
<td>4.50</td>
<td>rs13477272</td>
<td>3</td>
<td>81.63</td>
<td>98.28</td>
</tr>
</tbody>
</table>

Abbreviations: CFA: contextual fear acquisition, LOD: logarithm of the odds, Chr: chromosome, Mb: megabase, CFM: contextual fear memory, Ntg: non-transgenic
As background strain was also observed to significantly influence cognitive function in the Ntg-BXD population, we next performed genome-wide QTL mapping in the Ntg-BXD population to evaluate the specificity of our identified QTLs. No significant QTLs were identified as modifying CFA in the Ntg-BXD population, suggesting the locus on chromosome 3 identified in the AD-BXDs is a unique modifier of AD susceptibility. In contrast, multiple genome-wide significant QTLs appeared when considering Ntg-BXD CFM (Table 4-1, gene list available in Supplemental Table 10), including the region on chromosome 2 that was identifying as significantly regulating CFM in the AD-BXD population. This result suggests this region harbors variants that globally regulate cognitive decline across normal aging and AD. In support, when the AD and Ntg-BXD populations were mapped together, an overlapping region on chromosome 2 emerged as a significant regulator of population-level CFM (Table 4-1, gene list Supplemental Table 10). When mapping the AD and Ntg-BXD populations together, an additional unique locus was identified on chromosome 1 as a mediator of population-level CFA (Table 4-1, gene list Supplemental Table 11), again supporting our hypothesis that the locus on chromosome 3 interacts relatively specifically to modify AD progression and severity. Together, these results highlight multiple genomic regions that may be explored in future studies to slow the onset and progression of cognitive decline in both normal aging and AD.

**Inflammation and loss of synaptic genes underlie population-level cognitive decline in AD-BXDs**

Having identified multiple genetic loci that may contribute to cognitive decline across our population, we next sought to better understand transcriptional mechanisms that underlie AD-related cognitive decline. Specifically, we looked at transcriptional changes occurring from 6 to 14 months of age in the AD-BXDs by performing RNA sequencing on a hippocampal tissue from a total of 157 AD-BXD mice, 71 of which were 6 months of age (47 female/24 male) and 86 of which were 14 months of age (45 female/41 male). As a population, the AD-BXDs significantly decline in cognitive function from 6 months to 14 months of age (Figure 4-1), providing an ideal opportunity to understand changes that underlie the onset of disease throughout this period. Together, these mice represented a total of 80 sex/age groupings across 28 genetically diverse background strains, and biological replicates were averaged together for downstream analyses. Differential expression analysis using DESeq2 identified a total of 1278 genes that significantly change in expression (774 up-regulated and 504 down-regulated, adjusted p-value < 0.05) throughout the course of aging. Using GSEA, a slight but significant up-regulation of genes enriched for immune-related functions was observed (Figure 4-2A), suggesting neuroinflammation increases with age in the AD-BXDs. This increase seems to be driven by an increase both in microglia and astrocyte numbers, as up-regulated genes that show cell-type specificity in their expression profiles showed enrichment for both microglia and astrocyte localization (Figure 4-2B, hypergeometric test, p < 0.0001 and p = 0.02, respectively). This up-regulation of immune-related genes
Figure 4-1. Cognitive function declines with age across genetically diverse AD mice.

(A) At 6 months of age, heritable variation in CFA exists across AD-BXDs, as measured by freezing following the final shock during the training trial of CFC. (B) Variation in CFA also exists at 14 months. (C) As a population, the AD-BXD panel declines significant in CFA performance with age [ANOVA main effect of age, F(1, 354) = 3.3, p < 0.001], although the extent of this decline varies widely. (D) At 6 months of age, heritable variation in CFM exists across the AD-BXDs, as measured by total freezing during the 10-minute testing trial of CFC. (E) Variation also exists at 14 months. (F) As a population, the AD-BXD panel declines significantly with age [ANOVA effect of age, F(1, 354) = 3.5, p < 0.001], although as with CFA, the extent of this decline varies widely across background strain. *Raw data originally reported in (Neuner et al., 2018), replotted here. Each point in C and F represents a strain/age average.
Figure 4-2. Inflammation and loss of synapses likely underlie population-level cognitive decline in AD-BXDs.

(A) GSEA identified a significant enrichment for GO terms related to immune function among genes up-regulated with age across AD-BXDs. Designations before each term indicate category of GO term; MF: Molecular function, BP: Biological process, CC: Cellular compartment. (B) Genes up-regulated in 14 month-old AD-BXDs that show cell-type specificity in their expression profiles showed enrichment for localization to microglia (p < 0.001), astrocytes (p = 0.02), and myelinating oligodendrocytes (p < 0.001), but not oligodendrocyte precursor cells (OPCs), endothelial cells (EC), newly formed oligodendrocytes (New oligo.), or neurons. (C) Among genes down-regulated in 14 month-old AD-BXDS, GSEA identified a significant enrichment for terms related to neuron function and in particular, synapse localization. (D) Genes down-regulated in 14m AD-BXDs that show cell-type specificity in their expression profiles showed enrichment for localization to neurons (p < 0.001).
mirrors changes observed in brain tissue collected post-mortem from human AD patients relative to controls (International Genomics of Alzheimer's Disease, 2015; Zhang et al., 2013). Interestingly, there was also a significant enrichment among positively changed genes for expression in myelinating oligodendrocytes. In contrast, a robust down-regulation of genes specifically enriched for localization to synapses, with functions involved in channel activity was observed (Figure 4-2C). As expected based on GO term enrichments, down-regulated genes that show cell-type specificity in their expression profiles showed enrichment for neuronal localization (Figure 4-2D, hypergeometric test, p < 0.0001), reminiscent of neurodegeneration observed in the classical 5XFAD model (Oakley et al., 2006) and human AD patients (Crews and Masliah, 2010).

Characterization of pre-symptomatic AD transcriptional network

To begin to understand potential alterations in transcriptional networks that precede overt symptom onset, we first used WGCNA (Figure 4-3A) to identify networks present in the transcriptome of 6 month-old AD-BXD mice. Specifically, we used WGCNA to identify modules of highly correlated genes present among genes that were expressed at a level of at least one TPM in >20% of our samples. As groups of highly co-expressed genes presumably function in similar biological processes or pathways, module construction using only 6 month AD-BXD samples should provide a snapshot of the transcriptional profile within the pre-symptomatic population (Langfelder and Horvath, 2008). From WGCNA, we identified 43 modules ranging in size from 35 to 2038 genes. Of the 43 modules, 27 were enriched for at least one functional category (GO term) at FDR ≤ 0.05. A variety of GO terms were represented (Supplemental Table 12), including immune processes, neuronal function, mRNA binding, protein binding, and myelination, suggesting we are capturing a number of physiologically relevant processes occurring in young AD-BXD mice.

Select pre-symptomatic gene signatures correlate with cognitive function later in disease

Although as a population the AD-BXDs decline with age, the extent of cognitive decline varies widely across the panel. Background strain explains a large portion of observed variation in cognitive decline (19% of total variance in CFA and 15% of the total variation in CFM across 6 and 14 month-old AD-BXDs (Figure 4-1). To identify which, if any, of our identified modules correlate with this observed strain-specific variation in cognitive decline, we first summarized module expression using the module eigengene (ME) generated by WGCNA. In order to generate the ME, the expression of all genes in a given module is summarized via principal component analysis in order to obtain the first principal component, which explains the largest proportion of expression variation across a module. A ME is particularly useful as it reduces dimensionality of a module to a single representative measure, which then can be used to relate module expression to both genotypes and external phenotypes. Here, we correlate ME expression
Figure 4-3.  Identification of pre-symptomatic gene signatures predictive of AD resilience.
To profile gene networks present in pre-symptomatic AD-BXDs, we performed WGCNA on hippocampal RNA sequencing data. Modules were summarized by a single ME (i.e. the first principal component) to simplify network visualization and analyses (Langfelder and Horvath 2007). (A) Dendrogram depicting hierarchical clustering of all MEs; dissimilarity indices were calculated by subtracting the r value for any given correlation from 1. (B) Heatmap illustrating correlation strength across the ME network. (C) Expression of the 6 month pink module, as measured by the ME, significantly correlates with strain-matched 14 month CFA. (D) Expression of the 6 month darkgrey module significantly correlates with strain-matched 14 month CFM.
at 6 months to strain-matched cognitive performance at 14 months to obtain a quantitative measure of transcriptional networks that predict strain-specific susceptibility or resilience to AD-related cognitive decline. The module most highly correlated with CFA, as measured by freezing following the final shock on day 1, was the pink module (r = -0.4, nominal p = 0.02, Figure 4-3B), which was significantly enriched for over 200 GO terms at an FDR ≤ 0.05 (Supplemental Table 12). A vast majority of these terms involved immune system function, including immune system process, immune response, and regulation of cytokine production. Given the enrichment of immune-related terms among genes that increase with aging in the AD-BXDs (Figure 4-2A), the identification of an immune-enriched module as a critical predictor of late-life learning abilities suggests the ability of a strain to modify or fine-tune these processes as early as 6 months is critical for disease progression. A single 6 month module, the darkgrey module (Figure 4-3C) was identified as highly correlated with 14 month CFM performance as measured by the percentage of time spent freezing throughout the 10-minute test (r = 0.3, nominal p = 0.046). The darkgrey module was also significantly enriched for a number of GO processes, although a smaller number than the pink module (n = 50). A large number of these GO terms related to extracellular matrix organization, extracellular organelle and vesicle localization, and aspects of protein binding (integrin binding, heparin binding, calcium ion binding, etc.). Interestingly, the identification of the darkgrey module also highlights a putative role for brain vasculature in disease progression, as the module was significantly enriched in the GO terms blood vessel development, circulatory system development, cardiovascular system development, blood vessel morphogenesis, and angiogenesis (FDR ≤ 0.05).

Identification of specific drivers of module-trait associations

While GO enrichment analyses give a broad snapshot of specific functions and pathways that may be targeted to promote resilience, one of the strengths of network analysis is that specific ‘hub genes’, or drivers of intra-modular connectivity, can be identified. These genes represent ideal therapeutic targets with which to manipulate the transcriptome on a modular level, as they are highly connected to other genes within a given module. However, traditional hub gene identification in WGCNA relies simply on a measure of gene-gene correlation (Langfelder and Horvath, 2008), and sometimes identifies a number of indirect connections such that if gene A regulates gene B and gene B regulates gene C, gene A and C will be appear to be highly connected based on their mutual relationship with gene B. To limit the use of indirect connections in nominating hub genes, we utilized partial correlation analysis (Zhang et al., 2018), where measures of gene-gene correlations are obtained by conditioning on all other genes in the module, such that the only genes that would appear to be connected in the previous scenario are genes A & B and genes B & C, but not genes A and C. It was our hypothesis that this approach would nominate hub gene candidates that could be targeted to most efficiently and directly modulate the broader network. In addition to hub genes, we were interested in module members highly correlated with the specific trait of interest themselves, independent of relationship to the ME or other module members. In order to come up with a list of candidates that may be targeted to promote networks hypothesized to be
underlying resilience (i.e. the pink and darkgrey modules), we selected genes appearing both in the list of top 30 hub genes remaining after partial correlation analysis and which exhibited significant correlation with either CFA or memory or at a p-value ≤ 0.05. The top 4 candidates from each module were identified and their relationship with 14 month-old cognitive performance is highlighted in Figure 4-4. A number of candidates with known links to AD were identified, including low affinity immunoglobulin gamma Fc region receptor II-b (Fcgr2b) (Kam et al., 2013), as well as candidates not well studied in the context of aging or AD, including G-protein coupled estrogen receptor 1 (Gper1) (Briz et al., 2015), both validating our approach and identifying novel candidates that likely contribute to disease onset.

**Genomic region on chromosome two underlies variation in resilience networks**

As the AD-BXD population was derived from the BXD GRP, which has been densely genotyped, we can begin to investigate the contribution of specific genetic variants to observed phenotypes, including the variation in transcriptional networks described here. In order to identify regions of the genome involved in regulating observed variation in identified resilience networks, we performed QTL mapping using either the pink or darkgrey ME as a quantitative trait. No QTLs were identified for the pink module, but there was a significant QTL on chromosome 2 identified as a regulator of the darkgrey module (Figure 4-5A). Strains carrying a copy of the D allele at the peak QTL marker had lower expression of the darkgrey module, as represented by lower ME values (Figure 4-5B), and lower CFM performance at 14 months of age (Figure 4-5C). We confirmed higher ME values were associated with higher module member expression by relating the mean standardized expression level of all member genes (Mostafavi et al., 2018) to the WGCNA-derived eigengene (r = 0.93, p < 0.01). These results demonstrate the presence of genetic variants in the region that significantly modify the expression of the group of genes that make up the darkgrey module and contribute to variation in cognitive decline. The identified interval is large (~30 Mb), so pinpointing the causal variant is difficult. Three module members, sperm associated antigen 6 (Spag6), olfactomedin-like 2A (Olfm2a), and prostaglandin D2 synthase (Ptgds) appear in the QTL interval, although as they were not top hub genes identified via partial correlation analysis, it’s unclear if they play a causal role in regulating module expression. These positional candidates are highly correlated with hub genes Slc6a13 and Eln, suggesting they associate more strongly with the darkgrey CFM module than the pink CFA module. Exactly how genes and variants in the QTL regulate module expression remain to be elucidated; any number of positional candidates may play a role in transcription and gene regulation at the protein level rather than transcriptional level. Alternatively, a number of micro-RNAs (miRNAs) and long non-coding RNAs (lncRNAs) exist in the identified interval and may post-transcriptionally regulate module gene expression.
Figure 4-4. Identification of putative target genes that may be targeted to promote resilience networks.
A combination of hub gene identification via partial correlation analysis and single-gene correlation prioritization was used to identify likely candidates driving the association between 6 month modules and 14 month-old cognitive performance. (A) Gene expression of putative candidates Fcgr2b, Ctsh, Hcls1, and Itgb2 at 6 months is plotted against strain-matched 14 month-old CFA. (B) Gene expression of putative candidates Slc6a13, Eln, Acta2, and Gper1 at 6 months is plotted against strain-matched 14 month-old CFM.
Figure 4-5. Identification of genomic regions underlying variation in resilience networks. 
(A) A significant QTL was identified as regulating the expression of the darkgrey module (LOD = 3.8, 1.5 LOD interval = 15-44.5 Mb). (B) Strains carrying the D allele at the peak QTL marker exhibited significantly lower expression of the darkgrey module [t(1, 37) = 4.4, p < 0.001]. (C) Strains carrying the D allele at the peak QTL marker with strain-matched 14m cognitive data exhibited significantly worse CFM [t(1, 32) = 4.3, p = 0.0002].
Discussion

Utility of the AD-BXD for studying resilience

Resilience to AD, defined as better than expected cognitive functioning based on high pathology load or high-risk genetic background, has traditionally been difficult to study, both in mice and humans. Specifically, human individuals with intact cognitive functioning rarely enter the clinic, and if they do, lack of access to relevant brain tissue early in disease precludes the identification of causal molecular mediators of cognitive decline. Studies in the mouse have traditionally utilized only a single AD model (Onos et al., 2016), where genetic variance and phenotypic variance is low relative to a highly penetrant AD transgene. We overcome some of these barriers by utilizing the genetically diverse AD-BXD mouse population, as this series of genetically diverse inbred mice show highly variable susceptibility to disease (Neuner et al., 2019a). All of the AD-BXD mice harbor a high-risk genotype in the form of the 5XFAD mutation, however not all strains show the expected degree of cognitive decline that is typically seen as a result of the aggressive 5XFAD transgene (Oakley et al., 2006). Instead, a number of strains demonstrate a certain degree of cognitive resilience (Figure 4-1). Given the inbred nature of the panel, the AD-BXDs represent an ideal opportunity to study the molecular determinants of resilience, as we now have the opportunity to obtain brain tissue early in disease while repeatedly sampling genetically identical individuals later in life in order to phenotypically classify certain strains as either susceptible or resilient to disease onset and better understand which individuals would have gone on to develop severe AD dementia.

The role of modifiers genes in normal aging and AD

The creation of the AD-BXD panel enables genetic mapping to identify true modifier alleles that influence the onset and severity of disease. However, interpretation of these modifiers in the context of normal aging is complex, as the extent to which genetic mechanisms underlying normal cognitive aging and AD overlap is still unclear. In principle, modifier alleles of interest could act in one of two ways: 1) as general modifier alleles that contribute to a phenotype regardless of disease status, or 2) specific AD modifiers that exhibit some type of epistatic relationship with the 5XFAD transgene and show no effect on phenotype in Ntg-BXD mice. Interestingly, we see examples of both instances in our results, suggesting both mechanisms are at play in our panel. For example, the QTL for CFM on chromosome 2 showed a significant association with CFM abilities in both AD- and Ntg-BXD mice and as such, genes located in the chromosome 2 region likely represent broadly applicable therapeutic targets due to their role in both normal aging and AD. As an example of the second potential modifier mechanism, the QTL for CFA played no discernable role in regulating learning abilities in the Ntg-BXD population, suggesting this locus played an AD-specific role and may act as an epistatic modifier of the 5XFAD transgene. Genes located in this interval may help to inform underlying mechanisms that lead to the exacerbated cognitive decline observed
in AD and represent valuable therapeutic targets. These two examples highlight the particularly complex relationship between normal aging and AD. It is also interesting to note that, as in humans, our identified QTLs do not explain all of the variation in their respective traits that are attributable to genetic factors. While this can be improved by increasing mapping power, our results and the lack of a QTL in AAO mapping further illustrates the point that multiple small effect variants with different roles (e.g. additive, epistatic) in disease susceptibility contribute to a portion of the heritability of this complex disease.

**Hub genes as candidates for promoting resilience**

Here we identify two modules, the pink module and the darkgrey module, as gene networks present early in disease that correlate with cognitive function later in life. It is our hypothesis that the biological pathways and processes represented by these modules play a role in priming the brain for late-life susceptibility or resilience to AD. As such, we hypothesize that manipulating the expression of either of these modules in the desired direction (i.e. down-regulating the pink module or up-regulating the darkgrey module) early in life would maintain cognitive function late in disease, as observed in our resilient strains. Hub genes, or highly connected gene within a given module, represent viable targets to manipulate whole-module expression, and they likely influence expression of a number of their nearest neighbors. Here we nominate the top four candidates from each module as putative candidates that may be targeted to promote cognitive resilience. As the pink module was largely enriched for immune-related GO terms, a number of the candidates that emerged from this module related to immune function. Although neuroinflammation has repeatedly been identified as a point of therapeutic intervention in AD (Bronzuoli et al., 2016), our approach now provides precise targets that may be leveraged early in disease. One target, integrin subunit beta 2 (Itgb2) has been identified as a member of the module identified by Zhang and colleagues (2016) to be most highly associated with AD onset in human patients, suggesting that target in particular harbors translational relevance to human disease. Interestingly, the candidates nominated by prioritization of members in the darkgrey module have more varied roles, as suggested by the more diverse GO terms enriched within the darkgrey module. Two candidates, elastin (Eln) and actin, alpha 2, smooth muscle, aorta (Acta2) again highlight the critical role of brain vasculature in the maintenance of cognitive function, as Eln is a major structural component of arterial walls and Acta2 has been implicated in vascularization and vascular branching (Rudnicki et al., 1990; Wagenseil and Mecham, 2012; Yamada et al., 1999). The two remaining candidates, solute carrier family 6 member 13 (Slc6a13) and Gper1, are both ion channels and receptors that have been implicated in neuronal signaling and synaptic plasticity (Briz et al., 2015; Chazalon et al., 2018). Of particular interest from the darkgrey module is the gene Slc6a13, as this gene, a GABA transporter, has been identified as a potential blood biomarker in AD (Long et al., 2016), suggesting this gene (in addition to Itgb2) may have translational relevance to human disease. In summary, results here show that dampening specific immune system processes while enhancing vascularization and synaptic health early in life may promote cognitive
resilience later in disease. In combination with specific candidate gene prioritization, this analysis now provides specific candidates on which to focus future studies.

**Identification of genetic mediators of resilience**

Here we show that genetic background explains a large portion of variation (~15-20%) in cognitive functioning among AD-BXD strains. We then go on to identify a number of QTLs associated with strain-specific variation in behavioral and neuropathological variation (Table 4-1), as well as specific region on chromosome 2 that directly influences expression of one of our identified resilience modules. While many of the regions are too large in the mouse to narrow down specific gene variants involved in modifying resilience (as of yet), a bulk of evidence suggests this strategy is capable of identifying therapeutic targets relevant to AD in human populations as well. For example, portions of the identified QTL modifying the darkgrey module are syntenic to a region of chromosome 9 (specifically 9q34) in the human genome, and evidence that chromosome 9 may harbor variants that influence AD susceptibility has emerged from a number of linkage studies in FAD populations. Specifically, 9q34 was identified as linked to AD onset in autopsy-confirmed familial AD patients in 1999 (Kehoe et al., 1999) and again in a study using 466 FAD families in 2000 (Pericak-Vance et al., 2000). Future studies that add additional AD-BXD strains in order to increase mapping power may be able to narrow this genomic interval and identify causal gene variants, which would better inform human studies seeking to identify novel AD genes. As none of our identified hub genes reside in this QTL, it is likely that the genes located within the QTL exert their effect on module expression at the protein level, a possibility that will be examined in future studies. In addition, a number of non-protein coding genes exist in the interval, such as micro-RNAs and long non-coding RNAs. These types of molecules have been shown to exert large effects on transcriptional networks (Brown et al., 2014; Gamazon et al., 2012; Jalali et al., 2013), but were largely excluded from the WGCNA module assembly, as our RNA isolation and library preparation protocol targeted poly-A enriched mRNAs. Future work will examine the possibility that genetic variation in a number of these non-coding genes may work to influence transcriptional networks underlying AD resilience.
CHAPTER 5. CONCLUSIONS AND FUTURE DIRECTIONS

Utility of Genetic Diversity in Mouse Preclinical Studies

Overall, our results demonstrate the significant untapped potential of utilizing naturally occurring mouse genetic diversity to understand susceptibility to age-and AD-related cognitive decline. In particular, the BXDs represent a significant resource that enables the utilization of genetically diverse but inbred and reproducible mice in a well-controlled environment. This resource, used either alone or in combination with a well-established AD model like the 5XFAD model, allows for the identification and better understanding of specific genes and transcriptional networks that underlie variation in complex traits such as cognitive decline and neuropathology.

In addition to enabling a mechanistic understanding of complex traits, the use of genetic diversity in model systems is likely to greatly enhance the translational relevance of preclinical findings. The ultimate goal of preclinical studies focused on aging and AD is the eventual translation of identified candidates into viable human therapeutics or biomarkers of disease. Our results suggest the AD-BXD panel is a valuable resource to do just that, as we demonstrate high levels of overlap between the AD-BXDs and human AD at the genetic, transcriptional, and phenotypic levels. In addition, we validate CFC, particularly the acquisition phase, as a translationally valid task that is likely to share some of the same underlying mechanisms as current cognitive tests used in the human clinic. Together, the combination of new tools (i.e. AD-BXDs) and valid tasks (CFC), used at the appropriate time points, may enable, for the first time, the identification of genetic modifiers of AD susceptibility that can be targeted as new therapeutic opportunities. A number of viable candidates influencing AD progression were discussed in Chapter 4, although extensive validation studies will need to be carried out to confirm or refute the hypotheses presented here. While a significant endeavor, such studies will be greatly facilitated due to the rigor and reproducibility afforded by the AD-BXD panel. As there is currently no cure for AD, understanding the mechanisms that protect some individuals from developing this devastating disease would be instrumental in developing novel therapeutics and ultimately, finding a cure.

Finally, having a genetically diverse model of AD, especially one that is reproducible, greatly enhances our ability to find a therapeutic target likely to generalize to a wider population. For example, the majority of therapeutics tested in a single mouse model of AD failed to show significant results when tested on a genetically diverse population of human patients. Now, researchers can utilize AD-BXD strains showing a range of susceptibility to disease to first screen therapeutics and select those that have the most robust benefit across the population. In addition, the BXD panel also provides an important resource for pharmacogenomics, as the ability to screen more generalizable drugs across a panel of genetically diverse individuals would allow for more accurate estimates of both widespread utility of the therapeutic and potential unanticipated toxic side effects. Conversely, one could also screen drugs that have failed in human clinical trials across a range of AD-BXD strains and identify those that would have predicted the
drugs failure. These strains represent more accurate models of the human condition and may be utilized in future preclinical studies in addition to traditional mouse models of AD.

**Hp1bp3 as a Mediator of Normal Cognitive Aging**

In the context of normal aging, we have now demonstrated for the first time that \(Hp1bp3\) is a key modulator of cognitive decline. Specifically, we demonstrate \(Hp1bp3\) acts to modify cognitive abilities by altering the expression of a number of synapse-localized genes, leading to deficits in intrinsic hippocampal neuronal excitability and long-term potentiation (Neuner et al., 2018). Given 1) the robustness of our results across multiple genetic backgrounds, 2) the overlap of transcriptional changes observed following \(Hp1bp3\) KD and human aging (Figure 2-10), and 3) the misregulation of HP1BP3 observed in human patients (Figure 2-11), we speculate that treatments that restore \(Hp1bp3\) expression and/or function may improve cognition in patients with normal cognitive decline. These results further contribute to our understanding of both physiological functions of \(Hp1bp3\) and molecular mediators of memory formation and storage, as well as some of the hallmark symptoms of aging.

In addition to the selective regulation of gene expression, previous studies have implicated \(Hp1bp3\) in the regulation of chromatin structure (Dutta et al., 2014). Recent studies have demonstrated that long-term maintenance of chromatin structure is critical for maintenance of cognitive longevity (Lopez-Otin et al., 2013), providing an additional mechanism \(Hp1bp3\) may work through in order to influence cognitive aging. Future studies will address this possibility directly with targeted ATAC-seq (an assay for transposase accessible chromatin with high-throughput sequencing) in order to assess how chromatin accessibility changes with \(Hp1bp3\) KD. The discoveries made here, as well as in future studies, will help inform the development of therapeutics designed to target either \(Hp1bp3\) or its downstream effectors (such as those plasma membrane ion channels and receptors highlighted in Figure 2-4C and Figure 2-8B-C) in order to help promote cognitive longevity and reduce the risk of multiple age-related diseases.

**Limitations of the Current Study**

While the BXD panel incorporates a significant degree of genetic diversity, a number of genomic regions remain identical between the B6 and D2 mouse strains. Practically, this means that a number of genes are not perturbed by naturally occurring variation across the panel and thus, will not contribute to differences in trait variation. Functionally, this may mean that BXD strains are largely identical in a number of key pathways and physiological functions. However, the degree of genetic diversity seems to be sufficient for obtaining a wide range of cognitive performance and Aβ42 levels (Figure 2-1B and Figure 3-1C-E), suggesting key genes, pathways, and mechanisms involved in cognitive function and APP processing are perturbed by genetic variation across the BXDs. A putative way to overcome this potential limitation and incorporate
additional diversity would be to add additional inbred strains to our study design. This could be in the form of additional single inbred strains, such as in the Mouse Hybrid Diversity Panel (Lusis et al., 2016), or additional RI strains from other crosses (Churchill et al., 2004; Williams and Williams, 2017). In particular, the Collaborative Cross set of RI strains could be utilized to incorporate extreme amounts of genetic diversity, as it incorporates 8 progenitor genomes rather than the standard 2, including 3 wild-derived strains that together segregate for almost 50 million SNPs (Churchill et al., 2004).

One of the additional limitations of our study is the lack of consideration of tau pathology. Since the 5XFAD transgene is not generally thought to induce significant tau pathology (Oakley et al., 2006), we did not evaluate tau pathology across the AD-BXDs. As hyperphosphorylated tau is one of the main pathological events observed in AD, this is a general shortcoming of most currently available mouse models of AD. Naturally occurring variants in Mapt [see Sanger Mouse Genomes project (Keane et al., 2011)] segregate across the BXD panel and may influence the production of hyperphosphorylated tau (and ultimately, neurofibrillary tangles). As such, it is possible that some genetic backgrounds do exhibit tau pathology. These strains (if identified) would provide ideal models for preclinical screens in future studies and contribute to our understanding of the genetic variants modifying production and/or clearance of hyperphosphorylated tau, as tau and amyloid neuropathologies do not often occur together in currently used mouse models (Kitazawa et al., 2012). As tau pathology has been reported to be more strongly associated with cognitive function in AD than other pathologies (Brier et al., 2016), the AD-BXD panel may provide new opportunities to study the relationship between tau, amyloid, and cognitive function in the context of genetic diversity. Future studies should examine this possibility directly and evaluate levels of various tau species across the AD-BXDs. Alternatively, future studies could utilize a model that incorporates targeted MAPT mutations (Oddo et al., 2003; Saul et al., 2013) to produce a genetically diverse AD model with substantial tau pathology.

**Genetically Diverse Mouse Models as a Resource for Precision Medicine**

Given advances in genomic technologies, there has been an emphasis on developing precision medicine strategies (Collins and Varmus, 2015), or individualized therapeutics based on specific genetic or molecular variation. As numerous drugs designed to prevent or delay AD across the general population have failed, this strategy may prove useful for developing new therapeutics. Given the reproducible nature of the BXDs (and resulting F1 AD-BXDs), they provide an important resource for developing and validating precision medicine experimentally. As large numbers of genetically identical individuals can be easily generated, we have the opportunity to thoroughly investigate gene-by-treatment interactions across a population in a well-powered manner. In addition, as BXDs have been densely genotyped (Wang et al., 2016), and the parental B6 and D2 strains fully sequenced (Keane et al., 2011)], we have a wealth of genetic information available upon which to base predictions and develop individualized therapeutics. The substantial amount of data collected from each BXD strain, particularly the diverse transcriptional data, and made available at GeneNetwork.org also greatly
enhances the information from which we can draw upon to make predictions about the underlying drivers of disease susceptibility.

One of the key aspects of precision medicine involves understanding the genetic and/or molecular causes of the disease. The lack of detailed understanding of mechanisms underlying AD has led to a number of failed clinical trials and an overwhelming lack of disease-modifying treatments. One possible explanation is that the extreme heterogeneity of AD in terms of clinical, imaging, and pathological characteristics (Lam et al., 2013) is indicative of a number of distinct subgroups of AD, each with distinct etiologies. As such, studies aimed at identifying one unifying cause or mechanism and which group all subtypes together are unlikely to yield significant or meaningful results. In support of this idea, a recent GWAS identified a number of novel loci associated with distinct subtypes of AD as classified by performance on a variety of cognitive tasks (Mukherjee et al., 2018). A better understanding of each of these putative subgroups and underlying genetic and molecular causes is needed to facilitate precision medicine and development of individualized therapeutics. Given their translational relevance, our AD-BXD panel provides an unprecedented opportunity to use naturally occurring genetic variation in the mouse to better understand subtypes of AD and interrogate the presence or absence of specified comorbidities. Once identified, we can then utilize targeted treatments designed to delay or prevent the onset of disease in an individual mouse strain (or strains) displaying a defined subtype of AD. Being able to perform precision medicine in a well-defined, environmentally controlled preclinical setting is a foundational first step towards delivering individualized therapeutics in the clinic.

**Future Directions**

Here, we have made significant strides towards understanding genetic and molecular mechanisms that underlie susceptibility to age-related cognitive decline as well as cognitive and pathological symptoms of AD. The utility of both the BXDs and the 5XFAD-BXD RIB design is clearly demonstrated through the results reported here. However, the utility of this experimental design should not stop at cognitive aging and AD. Instead, this approach should serve as blueprint for researchers interested in other complex diseases, especially those that can be modeled using a dominant transgene or targeted mutation and are likely influenced by background genetic variation such as Amyotrophic Lateral Sclerosis (Lutz, 2018), Parkinson’s disease (Antony et al., 2011), and frontotemporal dementia (Roberson, 2012; Swarup et al., 2019), among others. As various genomic, transcriptomic, and bioinformatic resources designed to facilitate systems genetics analyses continue to improve, the utility of genetically diverse mice, particularly the BXDs and related crosses, is likely to continue to enhance our understanding of individual variation in disease susceptibility and yield life-changing therapeutics for the millions of individuals afflicted with these devastating diseases.
LIST OF REFERENCES


Bahrini, I., Song, J.H., Diez, D., and Hanayama, R. (2015). Neuronal exosomes facilitate synaptic pruning by up-regulating complement factors in microglia. Sci Rep 5, 7989. DOI: 10.1038/srep07989


reveal major transcriptional and tumor suppressor responses. Proc Natl Acad Sci U S A 101, 2173-2178. DOI: 10.1073/pnas.0308512100


Charlesworth, D., and Willis, J.H. (2009). The genetics of inbreeding depression. Nat Rev Genet 10, 783-796. DOI: 10.1038/nrg2664

dysfunction generates tonic inhibition in external globus pallidus neurons in parkinsonian rodents. Cell Rep 23, 1678-1690. DOI: 10.1016/j.celrep.2018.04.014


synaptic vesicle recycling efficiency and cognitive deficits in amphiphysin 1 knockout mice. Neuron 33, 789-804.


Kaczorowski, C.C., and Disterhoft, J.F. (2009). Memory deficits are associated with impaired ability to modulate neuronal excitability in middle-aged mice. Learn Mem 16, 362-366. DOI: 10.1101/lm.1365609


116


119


resource for systems genetics analyses of metabolic and cardiovascular traits. J Lipid Res 57, 925-942. DOI: 10.1194/jlr.R066944


125


128


Zhang, W., and Linden, D.J. (2003). The other side of the engram: Experience-driven changes in neuronal intrinsic excitability. Nat Rev Neurosci 4, 885-900. DOI: 10.1038/nrn1248


Figure A-1. Delivery of Hp1bp3-shRNA to the hippocampus does not affect cortical or cerebellar levels of HP1BP3.

(A) No effect of treatment on HP1BP3 levels in the cortex was observed on western blot (n = 3/grp, two-way ANOVA effect of strain F(1, 8) = 4.2, p = 0.08, effect of treatment F(1,8) = 0.02, p = 0.9, no interaction]. Right, band densities were adjusted for total protein loading using GAPDH levels and quantified. (B) No effect of treatment on HP1BP3 levels in the cerebellum was observed [n = 3/grp, effect of strain F(1,8) = 1.4, p = 0.3, effect of treatment F(1, 8) = 0.2, p = 0.6, no interaction]. Right, band densities were adjusted for total protein loading using GAPDH levels and quantified. Lysate from human 293T cells overexpressing HP1BP3 was used as a positive control.
Figure A-2. KD does not have non-specific effects on anxiety, activity, or body weight.

No effects of strain or treatment were observed on (A) anxiety as measured by the percentage of entries made into the open arms on EPM [effect of strain F(1, 40) = 1.8, p = 0.18, effect of treatment F(1,40) = 0.33], (B) activity as measured by total arms entered on the EPM [effect of strain F(1, 40) = 0.39, p = 0.54, effect of treatment F(1,40) = 0.00, p = 0.99], (C) activity as measured by the average speed in the EPM [effect of strain F(1, 40) = 0.46, p = 0.50, effect of treatment F(1,40) = 0.06, p = 0.82], or (D) body weight [effect of strain F(1, 41) = 0.05, p = 0.83, effect of treatment F(1,41) = 3.0, p = 0.09]
<table>
<thead>
<tr>
<th>Upstream Regulator</th>
<th>Molecule Type</th>
<th>Predicted Activation State</th>
<th>Activation Z-score</th>
<th>P-value of Overlap</th>
<th>Target Molecule in Dataset</th>
<th>Corresponding miRNA in Sequencing Data</th>
<th>Log2 Fold Change</th>
<th>Adjusted P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mir-21</td>
<td>microRNA</td>
<td>Inhibited</td>
<td>-4.94</td>
<td>4.35e-13</td>
<td>AIF1, A0AH, C1R, Casp12, CASP4, CCL3L3, CD180, CD274, CDKN1A, CLEC5A, CSF3R, CXCL10, DDAH2, DX58, FCGR1A, FCGR2A, FGL2, GBP2, GBP5, GBP6, GBP7, Gbp8, Gvin1 (includes others), Gzmb, IFI16, Ifi47, Ifi51, IRF1, IRGM, Irgm1, ITK, Klrk1, LTB, MMP2, MSH2, NLRC5, Nlrp1a, NLRP3, OAS2, OAS3, PDCD4, PECAM1, PER2, PER3, PILRB, PSME2, SIGLEC1, SLAMF1, Sf1n2, Sp100, SPRY2, STAT1, STAT2, SUMO1, TAP1, TGBF1, TGM1, Tgpt1/Tgpt2, TIAM1, TLR1, TLR2, TLR4, TNF, TRAF1, TREM2 AMPD3, APOBEC1, BST1, C15orf48, C5AR1, CAMKK1, CD180, CD274, Cd33, CD93, Clec2d (includes others), CLEC4E, CSF1, EMP1, FAM20C, FGL2, GPR65, GSTM5, HLA-DOB, HPGDs, IFI16, Ifitm6, IL1A, IL1B, IL1RN, INPP5J, Ly6a (includes others), LY96, MLKL, MMP13, MMP25, Mt2, NLRP3, PURB, RIPK1, RIPK3, S100A6, SCARB1, Serpina3g (includes others), SGLPL1, STAT3, TGFB1, Tlr13, TLR4, TLR7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mir-223</td>
<td>microRNA</td>
<td>Activated</td>
<td>3.04</td>
<td>4.56e-7</td>
<td></td>
<td>mmu.miR.223.3p</td>
<td>1.15</td>
<td>0.09</td>
</tr>
<tr>
<td>miR-124-3p (miRNAs w/seed AAGGCAC)</td>
<td>Mature microRNA</td>
<td></td>
<td>-1.41</td>
<td>0.0008</td>
<td>Ccl2, CTDSP1, EGR1, HDAC5, ITGB1, LAMC1, LMNB1, MAPK14, PTBP1, SYCP1</td>
<td>mmu.miR.1247.5p</td>
<td>-0.16</td>
<td>0.51</td>
</tr>
</tbody>
</table>
Table A-1. Continued.

<table>
<thead>
<tr>
<th>Upstream Regulator</th>
<th>Molecule Type</th>
<th>Predicted Activation State</th>
<th>Activation Z-score</th>
<th>P-value of Overlap</th>
<th>Target Molecule in Dataset</th>
<th>Corresponding miRNA in Sequencing Data</th>
<th>Log2 Fold Change</th>
<th>Adjusted P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mir-146</td>
<td>microRNA</td>
<td>Inhibited</td>
<td>-2.74</td>
<td>0.002</td>
<td>DUSP1, ERBB4, FADD, IL1B, MMP2, MYC, MYD88, PTGS2, RUNX1, SOCS1, STAT1, TNF, WASF2</td>
<td>mmu.miR.146a.5p</td>
<td>1.23</td>
<td>0.07</td>
</tr>
<tr>
<td>Mir-155</td>
<td>microRNA</td>
<td></td>
<td>-1.52</td>
<td>0.002</td>
<td>CCL2, CCL5, CD68, CD69, CXCL10, DNNJA2, DNNJB1, DYNCI11, FADD, HK2, HMOX1, IL12B, IL1B, INPP5D, MMP3, MSH2, MYD88, PTGS2, RIPK1, SOCS1, SOCS3, STAT3, TNF, TP53, TRIM32, WEE1</td>
<td>mmu.miR.155.5p</td>
<td>2.07</td>
<td>0.07</td>
</tr>
<tr>
<td>miR-155-5p (miRNAs w/seed UAAUGCU)</td>
<td>Mature microRNA</td>
<td>Inhibited</td>
<td>-3.28</td>
<td>0.01</td>
<td>MAPKAPK2, MKI67, SOCS3, YAP1</td>
<td>mmu.miR.155.5p</td>
<td>2.07</td>
<td>0.07</td>
</tr>
<tr>
<td>miR-483-3p (miRNAs w/seed CACUCCU)</td>
<td>Mature microRNA</td>
<td></td>
<td>-1.98</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mir-103</td>
<td>microRNA</td>
<td></td>
<td>-0.64</td>
<td>0.02</td>
<td>ADAM10, AIP, ARNT, CAV1, HIF1A, TP53</td>
<td>mmu.miR.103.3p</td>
<td>-0.19</td>
<td>0.21</td>
</tr>
<tr>
<td>miR-125b-2-3p (miRNAs w/seed CAAGUCA)</td>
<td>Mature microRNA</td>
<td></td>
<td>0.02</td>
<td></td>
<td>BAK1, BMF, TP53</td>
<td>mmu.miR.125b.2.3p</td>
<td>0.10</td>
<td>0.63</td>
</tr>
<tr>
<td>Mir-15</td>
<td>microRNA</td>
<td></td>
<td>0.31</td>
<td>0.02</td>
<td>BACE1, Ccl2, CD82, EIF4G2, FGF1, HSPA1A/HSPA1B, IL1B, MAPK3, PIM1, PTGS2, TIMP1, TP53, UCP2, ABCA1, CDKN1A, DUSP2, FAS, FGF9, MAPK9, NPS3, RB1, RUNX1, STAT3, THBS1, ZBTB7A</td>
<td>mmu.miR.15a.5p</td>
<td>0.19</td>
<td>0.57</td>
</tr>
<tr>
<td>Mir-17</td>
<td>microRNA</td>
<td></td>
<td>-1.60</td>
<td>0.02</td>
<td></td>
<td>mmu.miR.17.3p</td>
<td>0.51</td>
<td>NA</td>
</tr>
<tr>
<td>Upstream Regulator</td>
<td>Molecule Type</td>
<td>Predicted Activation State</td>
<td>Activation Z-score</td>
<td>P-value of Overlap</td>
<td>Target Molecule in Dataset</td>
<td>Corresponding miRNA in Sequencing Data</td>
<td>Log2 Fold Change</td>
<td>Adjusted P-value</td>
</tr>
<tr>
<td>--------------------</td>
<td>---------------</td>
<td>----------------------------</td>
<td>---------------------</td>
<td>-------------------</td>
<td>-----------------------------</td>
<td>--------------------------------------</td>
<td>----------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Mir-132</td>
<td>Mature microRNA</td>
<td></td>
<td>0.03</td>
<td>0.03</td>
<td>BDNF, CCL2, IRAK4, NR4A2</td>
<td>mmu.miR.132.3p</td>
<td>0.67</td>
<td>0.10</td>
</tr>
<tr>
<td>miR-155-3p</td>
<td>Mature microRNA</td>
<td></td>
<td>0.03</td>
<td>0.03</td>
<td>DNAJA2, DNAJB1</td>
<td>mmu.miR.155.3p</td>
<td>1.48</td>
<td>NA</td>
</tr>
<tr>
<td>miR-324-5p</td>
<td>Mature microRNA</td>
<td></td>
<td>0.04</td>
<td></td>
<td>RUNX1, SRF</td>
<td>mmu.miR.324.5p</td>
<td>-0.16</td>
<td>0.51</td>
</tr>
<tr>
<td>miR-125b-5p</td>
<td>Mature microRNA</td>
<td></td>
<td>0.71</td>
<td>0.04</td>
<td>AJUBA, APLN, BAK1, ENTPD4, MAP2K7, MAPK14, MYD88, RABL6, TNF, ZNF385A</td>
<td>mmu.miR.125b.5p</td>
<td>0.06</td>
<td>0.79</td>
</tr>
<tr>
<td>Mir-10</td>
<td>microRNA</td>
<td>Inhibited</td>
<td>-2.26</td>
<td>0.04</td>
<td>BAK1, BMF, HOC, MYD88, PRDM1, PTPN18, PTPN7, SUV39H1, TNF, TP53 ADAM17, ANXA11, DUSP2, EGLN3, ENTPD4, Foxp1, GPX7, MECP2, NDRG3, NFATC1, OSMR, PALM, RAB6B, TMED3, TRIB1 GFAP, KDR, MMP2, PDCD4, PECAM1, SPRY2, TGFB2</td>
<td>mmu.miR.10a.3p</td>
<td>1.01</td>
<td>NA</td>
</tr>
<tr>
<td>miR-122-5p</td>
<td>Mature microRNA</td>
<td></td>
<td>-0.78</td>
<td>0.04</td>
<td>ADAM17, ANXA11, DUSP2, EGLN3, ENTPD4, Foxp1, GPX7, MECP2, NDRG3, NFATC1, OSMR, PALM, RAB6B, TMED3, TRIB1 GFAP, KDR, MMP2, PDCD4, PECAM1, SPRY2, TGFB2</td>
<td>mmu.miR.122.5p</td>
<td>-0.08</td>
<td>NA</td>
</tr>
<tr>
<td>miR-21-5p</td>
<td>Mature microRNA</td>
<td></td>
<td>-1.56</td>
<td>0.05</td>
<td>BDNF, CCL2, IRAK4, NR4A2</td>
<td>mmu.miR.21.5p</td>
<td>0.86</td>
<td>NA</td>
</tr>
</tbody>
</table>
Figure B-1. Cognitive traits are not directly associated with a variety of peripheral traits.

(A) AAO of working memory deficits, as measured in the y-maze, is not correlated to (left) the average distance traveled or (right) average number of total arms entered. (B) Neither (left) CFA or (right) CFM is significantly correlated to pain sensitivity as measured by post-shock reactivity. (C) Neither CFA nor (D) CFM is correlated to sensorimotor abilities at either (left) 6 months or (right) 14 months of age. (E) Neither CFA nor (F) CFM is correlated to anxiety, as measured by percent time spent in open arms on the EPM, at either (left) 6 months or (right) 14 months of age.
Figure B-2. AD-BXD mice develop plaques throughout the hippocampus and cortex.
The percentage of (A) the hippocampus and (B) the cortex covered by amyloid plaques, as detected by Aβ42 IHC and ImageJ particle analysis in a subset of AD-BXD mice, highly correlates with the overall levels of Aβ42 as detected by ELISA in a subset of mice where both assays were performed (n = 37). (C) Representative images of plaque deposition in AD-BXD mice. (D) No staining was detected in 6 month-old Ntg-BXD mice (n = 3). Scale bar = 1 mm.
Figure B-3. Cognitive traits are not directly associated with Aβ42 levels. Aβ42 levels, as measured by ELISA at 6 months, are not significantly correlated to (A) 6 month CFA, (B) 6 month CFM, or (C) AAO of working memory deficits. Aβ42 levels, as measured by ELISA at 14 months, are not significantly correlated with (D) 14 month CFA, (E) 14 month CFM, or (F) AAO of working memory deficits.
Figure B-4. Background strain does not significantly modify 5XFAD transgene expression or endogenous App/Psen1 levels.
Expression of mutated human APP/PSEN1 and endogenous mouse App/Psen1 evaluated in a larger subset of mice [n = 293 (177 females/116 males across 28 strains)]. (A, left) AD-BXD mice exhibited significantly greater hAPP expression [t(1, 291) = 92.3, p < 0.001]. Across the AD-BXDs, there was no significant effects of age, sex, or (A, right) background strain. All 28 strains are shown here in comparison to only strains with Aβ42 data in Figure 1. (B, left) AD-BXD mice exhibited significantly greater hPSEN1 expression [t(1, 291) = 107.6, p < 0.001]. Across the AD-BXD panel, there were no significant effects of age, sex, or (B, right) background strain on expression. (C, left) AD-BXD mice exhibit a slight but statistically significant decrease in mApp expression relative to Ntg-BXDs [t(1,291) = 2.6, p = 0.01]. However, across the AD-BXD panel, there was no significant effect of age, sex, or (C, right) background strain on mApp expression. (D, left) AD and Ntg-BXDs exhibit comparable endogenous Psen1 expression. Across the population, there were no main effects of 5XFAD genotype, age, sex, or (D, right) AD-BXD background strain on Psen1 expression.
Figure B-5.  *Apoe* genotype effects CFM, but not CFA, in Ntg-BXD mice.
(A) Across Ntg-BXD mice, there was no effect of *Apoe* genotype on CFA [F(1, 280) = 0.3, p = 0.6]. There were no additional effects of sex [F(1,280) = 1.5, p = 0.2], age [F(1, 280) = 0.2, p = 0.7], or interactions with *Apoe* genotype. (B) Across Ntg-BXD mice, those carrying one copy of the D allele at *Apoe* performed significantly worse on CFM tests than mice with two copies of the B allele [F(1, 280) = 12.0, p = 0.001]. There were no additional effects of sex [F(1, 280) = 0.3, p = 0.6], age [F(1, 280) = 2.3, p = 0.13], or interactions with *Apoe* genotype.
Figure B-6. AD genetic risk score defined in Ntg-BXD mice shows no relationship with cognitive outcomes.

(A) Ntg-BXD strains were stratified into impaired (below population average) and unimpaired (above population average) based 6 month CFM performance. (B) GRS were calculated for each strain based on allelic composition of 21 genes known to confer risk for AD, with the risk allele for each gene defined as that which appeared more frequently in the impaired population shown in (A). (C) GRS showed no relationship to 14 month CFA across Ntg-BXDs, (D) CFA in 14 month-old AD-BXD mice, or (E) Aβ42 load at 14 months in AD-BXDs.
Figure B-7.  AD-BXD panel exhibits high transcriptional concordance with human AD.

Hargis and colleagues (Hargis and Blalock, 2017) recently identified a transcriptional signature of AD, consisting of the top 10% of up-regulated and down-regulated genes common to four different human AD datasets. To evaluate transcriptional concordance between our panel and human AD, the direction and magnitude of change was compared across mouse and human datasets. Specifically, the log₂ fold change (FC) of gene expression between AD- and Ntg-BXDs is plotted on the y-axis, while the log₂FC of gene expression between human AD patients and controls plotted on the x-axis. (A) Log₂FC of 39/60 (65%) AD signature genes is concordant between our mouse panel and a study by Blalock et. al., 2011 (Blalock et al., 2011). (B) Log₂FC of 39/59 (68%, one AD gene not detected in human dataset) AD signature genes is concordant between our mouse panel and a study by Blalock et al., 2004 (Blalock et al., 2004). (C) Log₂FC of 39/60 (65%) AD signature genes are concordant between mouse and human using dataset by Hokama et al., 2014 (Hokama et al., 2014). Secondary analyses of human datasets, see (Hargis and Blalock, 2017). Each point represents a gene; log₂FC with opposite directionality have been highlighted in red.
VITA

Sarah Marie Neuner was born in Voorhees, NJ in 1991 and grew up in Galloway, NJ. She completed her undergraduate studies at Richard Stockton College of NJ (now Stockton University) in Galloway, NJ and received a Bachelor of Science in Biology with a minor in Chemistry in the fall of 2012. She was also a part of the honors program at Stockton College. She was accepted into the University of Tennessee Health Science Center Biomedical Sciences Program as a member of the Neuroscience track in August 2013. She joined the lab of Dr. Catherine Kaczorowski in November of 2013 and received an F31 Individual Predoctoral Fellowship from the National Institutes of Health in under the co-sponsorship of Dr. Robert Williams in 2015. In 2016, the Kaczorowski lab relocated to The Jackson Laboratory (JAX) in Bar Harbor, ME and Sarah moved to JAX to complete her doctoral research. Dr. Williams served as Sarah’s co-advisor throughout her graduate career and remains her official mentor at UTHSC. During her time in graduate school, Sarah attended a number of local, national, and international meetings and gave a number of poster and oral presentations. She was also awarded several travel scholarships and assisted in several mentoring, outreach, and leadership programs at both UTHSC and The Jackson Laboratory. Sarah will graduate from UTHSC and receive her Ph.D. with a major in Biomedical Sciences and a concentration in Neuroscience in May of 2019.

Publications


