SPECIFIC AIMS

Alzheimer's disease affects over 35 million people worldwide, and there is no known cure. Although some treatments are available that can slow the disease, little progress has been made in understanding or preventing late-onset AD (LOAD, or sporadic AD), the most common form of the disease. Despite decades of research, the failure of recent drug trials strongly suggests new approaches are essential to develop treatments for LOAD. One of the major hurdles to developing new treatments is a lack of understanding of the fundamental genetic pathways that increase Alzheimer's disease susceptibility, initiation and propagation. To address this, we propose a novel systems genetic strategy that utilizes the strengths of The Jackson Laboratory.

Recent advances in mouse genetic resources combined with innovative computational methods make the laboratory mouse an increasingly useful model for understanding the biology that underlies Alzheimer's disease. LOAD is likely a polygenic disorder with causal perturbations in multiple genes. Therefore, we propose to use transcriptional profiling, computational analyses and functional testing to model the interactions between the top ten candidate LOAD genes. We will classify these LOAD genes according to the transcripts and functional processes they affect. The transcript data will be used to select pairs of LOAD genes for double gene deletions and apply novel computational methods to order genes into pathways/networks that regulate processes involved in LOAD. We will use these pathway models to construct new mouse models of LOAD and prioritize and validate potential targets for therapeutic intervention. We have three specific aims:

Specific Aim 1: To identify biological processes impacted by heterozygous and homozygous deletions of the ten strongest candidates genes for LOAD using RNAseq. We will perform genome-wide transcriptional profiling of the cortex and hippocampus in mice carrying mutations in the top 10 LOAD genes (as ranked by Alzgene). We will profile C57BL/6J mice carrying the human *APOE4* allele (the greatest risk factor for AD) and use this mouse as a reference strain for all further studies. Using APOE4 as a sensitizer will maximize the likelihood of identifying AD-relevant pathways. For all other LOAD genes, we will (where possible) separately test both heterozygous and homozygous deletions on the *APOE4* background to identify differential effects between gene dosage and gene deletion. Although the specific variations in the human orthologs are not known, our approach will allow us to identify many of the same transcripts/pathways that are dysregulated by variations in the human LOAD genes.

Specific Aim 2: To map complex genetic pathways relevant to LOAD with computational modeling of double-knockouts. We will apply a novel computational strategy combined with functional testing to infer genetic models of how variations in candidate genes interact positively (enhanced function) or negatively (suppressed function) in their contributions to Alzheimer's disease. First, singular value decomposition of transcriptome profiling data (Aim 1) will be used to computationally predict which pairs of LOAD genes are the best candidates for interactions and which biological processes they cooperatively affect. Second, transcript profiles from these best candidate pairs (generated using double-mutant mice) will be used to determine whether the genetic interactions are positive or negative. Finally, combined analysis of pleiotropy and epistasis (CAPE) will be applied to all single and pair-wise transcriptome data to produce a quantitative network model for Alzheimer's disease. Despite not knowing the specific variations in the human orthologs, CAPE will be able to predict the effects of both gain and loss of function of LOAD genes.

Specific Aim 3: To generate mouse models for LOAD and perform early validation of novel therapeutic targets. Mice mutant for multiple LOAD genes will be generated and assessed for Alzheimer-relevant phenotypes. In addition, novel therapeutic targets (predicted from the computational modeling in Aim 2) will be tested genetically for their ability to inhibit disease phenotypes. First, we will assess whether hAPOE4 mice mutant for *CLU*, *BIN1*, and *ABCA7* exhibit AD phenotypes as suggested by existing genetic knowledgebases. Second, additional mutant combinations (inferred from CAPE) will be tested for AD phenotypes. These genetic models will identify genes with robust disease-promoting or disease-protective effects on LOAD-associated biological process. Thirdly, disease-protecting candidate genes will be functionally tested as new therapeutic targets for AD by using genetic and pharmacological approaches.

RESEARCH STRATEGY

A. Significance

A multi-disciplinary approach to decipher the genetics of Alzheimer's disease
Alzheimer's disease (AD) is the leading cause of age-related dementia affecting more than 35 million people worldwide [1,2]. No treatments are available that prevent AD and only a few treatments are able to slow the progression of the disease. Recent drug trials that held great promise for new treatments have failed and the pharmaceutical industry is beginning to focus their efforts in other directions. The development of new treatments is limited by a lack of basic understanding of the biological processes dysregulated in AD. To determine the critical processes involved in AD it is essential to develop a more complete understanding of how genetic factors increase risk for late-onset AD (LOAD). To this end, we propose a systematic investigation into the biological pathways and gene networks relevant to AD in genetically tractable mouse models. This effort will be collaboratively led by a molecular geneticist (Howell, JAX) and a computational biologist (Carter, JAX) with expert input from a behavioral neuroscientist (Chesler, JAX) and a human geneticist (see letter of support, Prof. Julie Williams). We will use transcriptional profiling (Aim 1) combined with iterative computational modeling (Aim 2) to construct and refine networks of genes that underlie the pathology of AD. These networks will be functionally tested and culminate in the generation of new models for LOAD and the identification of novel drug targets (Aim 3).

Addressing multiple genes and pathways in a polygenic model of Alzheimer's disease
AD is a complex disorder with both genetic and environmental factors determining multiple disease outcomes (susceptibility, initiation, rate of progression and severity). Genetic linkage and genome-wide association studies (GWAS) have associated multiple genes with LOAD [4-8]. Inheritance of the e4 allele of the *APOE* gene provides the greatest known risk for AD and recently, variations in *TREM2* have been shown to confer a similar increase in risk of developing AD [9-11]. Meta-analysis of all GWAS (from Alzgene [12]) has ranked hundreds of genes to provide the best candidates for our study. Due to the diverse mouse resources at our disposal, we will be able to assess at least nine of the top ten genes including *APOE4*, *TREM2*, *BIN1*, *CLU* and *ABCA7*. As mouse resources continue to develop, we anticipate being able to incorporate additional genes (including novel findings) into our study. To determine how candidate genes interact to cause LOAD, we will critically test how multiple genes combine to affect known and novel AD-relevant phenotypes. By systematically assessing which genes/pathways are impacted by perturbations in single genes, then pairs of genes and ultimately sets of genes, we will construct a detailed picture of the fundamental biological processes that are altered to cause AD.

A modeling strategy to determine the Alzheimer's disease-relevant role(s) of LOAD genes
In most cases, the causative human sequence variations in LOAD genes are not known. We have devised a strategy that circumvents this limitation by using genetic perturbations to determine the AD-relevant functions of these genes rather than directly modeling the effects of the exact human variation (Figure 1). LOAD genes are reported to function in a variety of different processes including immune modulation and Aβ clearance. However, this is based on incomplete knowledge of both the genes themselves and the known events occurring in AD. For a more complete understanding of how these genes associated with LOAD impact the disease, we will measure the effects of manipulating them in mouse. Irrespective of the anatomical differences between humans and mice, greater than 90% of genes associated with LOAD have a single ortholog in mice (www.ensembl.org). Therefore, the genetic networks in which these genes function are expected to be largely conserved. Although previous studies using mouse models have had limited success, our strategy of combining physiological and genomic data with advanced bioinformatic modeling is specifically designed to uncover functionally relevant interactions affecting multiple biological processes that underlie LOAD.

Novel computational methods are necessary for dissecting complex traits
High-throughput technologies enable high-resolution genotyping and multi-dimensional phenotypic characterization, with the ultimate goal of providing precise genetic models of complex biological systems. The success of this approach is contingent upon the development of analytical methods to dissect genetic complexity and translate large-scale data into specific, testable hypotheses. The proposed research directly addresses this need through the application of recently developed methods to understand the genetics of age-

related diseases in a complex mouse model system. This proposal is significant in two areas: it will apply the Combined Analysis of Pleiotropy and Epistasis (CAPE) technique for use in mammalian systems; and it will concurrently use these methods to model and validate the biological processes that underlie LOAD in a mouse model system. The primary novelty of the CAPE method is to simultaneously analyze transcriptional profiles across many combinations of genetic perturbations, thereby identifying the best genetic model using all available information [13]. These new methods were designed for complex traits in model systems. However, they have not previously been applied to study age-related, genetically complex diseases such as AD. This proposal addresses a genetic model of complex disease using a study design of single and double-knockouts assayed across multidimensional phenotypes, providing an ideal application for the CAPE method. At the same time, CAPE is ideally suited to analyze how multiple candidate genes combine to affect different aspects of LOAD biology.

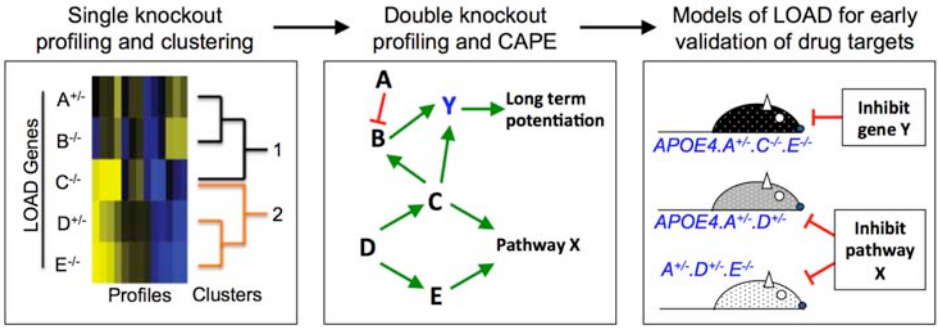

Figure 1: Our pipeline to determine new models of LOAD for early validation of drug targets (using hypothetical LOAD genes (A-E) and simulated data). Single-knockout mice will be analyzed to determine double-knockout combinations to test. CAPE computational analysis will be performed to obtain networks of genetic interactions that affect specific biological pathways and processes, from which we will predict and test novel mouse models of LOAD.

B. Innovation

Developing extensive mouse genetic resources to aid Alzheimer's disease research
Animal models of complex human diseases such as LOAD will require advances in genetic analysis. Previous attempts to model AD in mice have not been completely successful. This is likely in part due to a lack of understanding of the complex genetic interactions that affect LOAD-relevant biology. Here, we propose a novel pipeline for the derivation of mouse models for complex traits. Our pipeline uses transcriptional profiling data from two separate AD-relevant brain regions to construct networks describing the effects of mutations in LOAD genes (Figure 1). These networks directly inform which combinations of mutant genes will provide new mouse models of AD.

Extending the Alzheimer's knowledgebase in *GeneWeaver*, an online platform for genetic research
Research studies increasingly generate disparate data types that are distributed across many data portals, making integrative analysis of genetic systems a laborious and inefficient process. To ensure access to our data and, importantly, the ability to readily integrate our results with published findings, we will deposit all data in *GeneWeaver* [13]. *GeneWeaver* is an online system to query functional genomics data that bridges species, tissues, and experimental platforms, with a current focus on neurological phenotypes. This resource was created and is maintained by Dr. Elissa Chesler, our Co-Investigator. Timely integration of our data will allow the broader research community to efficiently use our data in their own studies of Alzheimer's disease and other neurodegenerative conditions.

Iterative computational modeling to construct predictive networks for Alzheimer's disease
As part of our pipeline, we will adapt and apply our strategy developed in simpler organisms ([14], see preliminary data) to dissect genetic interactions and pleiotropic effects of multiple LOAD genes on the biological processes that underlie the disease phenotypes. The quantitative basis for the analysis enables specific, testable predictions to be made for novel perturbations. Instances of model failure can then be used to direct additional rounds of data acquisition and model refinement. This is the first time this approach has been applied to a genetically complex age-related disease. The method will also serve as a template for data-driven genetic modeling of any age-related complex disease.

Innovative method to prioritize therapeutic targets for early validation
Failures in the ability to translate basic research findings into therapeutic targets for AD is due to the lack of testable systems that accurately model the human disease. Existing mouse models recapitulate some but not all components of early-onset AD. No mouse models exist for LOAD. Our approach provides new mouse models and a deeper understanding of the underlying processes that contribute to LOAD. Our network modeling approach allows a more precise identification of novel therapeutic targets. By testing these targets in mouse models that more comprehensively recapitulate AD-relevant processes, we will be able to better predict the efficacy of candidate therapeutic agents.

C. Approach

Preliminary Data

Transcriptional profiling by RNA-seq identifies pathways relevant to Alzheimer's disease
We propose to use transcriptional profiling to measure AD-relevant changes in the brain. To verify that transcriptional signals are associated with AD, we have used similar methodology to profile brains of existing AD models (e.g. C57BL/6J.*APP*^{swe}.*Psen1*^{de9})[15] and compared transcript profiles to control mice. Sixteen mutants (at 4-6 mos, early stages of AD in this mouse strain) were compared to eight age- and sex-matched controls. Approximately 6,000 genes with significant differential expression were detected (q<0.001). Pathways enrichment analysis (using DAVID) identified the Alzheimer's disease pathway (KEGG ID:mmu05010) as significantly overrepresented in our gene list (see Figure 2).

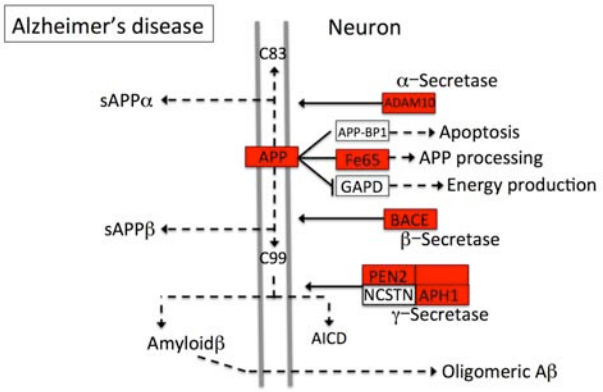

Figure 2. Alzheimer's-relevant genes are differentially expressed in the brain of AD mouse models. The figure shows a section of the KEGG pathway relating to Alzheimer's disease. In total **71** of the **183** genes in the pathway are differentially expressed (p<0.001).

Assessing Alzheimer's disease-relevant phenotypes in mice.
In this proposal, we will use a battery of methods to assess Alzheimer's-relevant phenotypes in mice. These include, but are not limited to, neuronal cell dysfunction, plaque deposition, synapse loss, astrogliosis and microgliosis. For example, β−amyloid build up can be assessed using ELISA (Figure 3A) and/or visualization by immunofluorescence with antibodies such as 4G8 (Figure 3B), 6A10 or Aβ42 directly (not shown). Injury/damage to neuronal axons can be assessed using phosphorylated neurofilament (pNFL – Figure 3C). Finally, neuronal cell dysfunction and loss can be assessed using antibodies that recognize neurons (such as NeuN), and counting NeuN-positive cells in different brain regions.

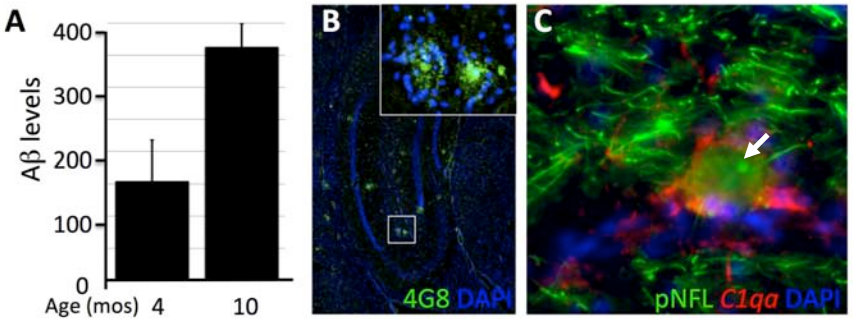

Figure 3: Assessment of AD phenotypes in mice. We will use a variety of methods to visualize AD phenotypes including (**A**) AB levels by ELISA, (**B**) plaque deposition, (**C**) and neuronal dysfunction by immuno-fluorescence. As a major goal is identify mouse models with frank neuro-degeneration, particular care will be taken to identify neuronal dysfunction. In the example shown, an antibody to phosphorylated neurofilament (pNFL) marks axonal swellings (or dystrophic neuritis, arrow) present in neurons in the hippocampus. Microglia are labeled using a riboprobe to *C1qa*.

Pathway/network analysis identifies gene networks for Alzheimer's disease

We have used pathway analysis to construct gene networks relevant to AD. Ingenuity Pathway analysis (IPA) Knowledgebase (http://www.ingenuity.com) is a repository of biological interactions and functional annotations created from individually modeled relationships between proteins, genes and other biologically relevant entities. We used the 48 genes that are considered significantly associated with Alzheimer's disease (based on Alzgene meta-analysis [12]) as seeds (Figure 4). This network incorporates nine of the top 10 ranked genes. Based on this analysis, *APOE*, *ABCA7*, *CLU* and *BIN1* are considered network "hubs" and have been targeted for mutation analysis (Aim 3a).

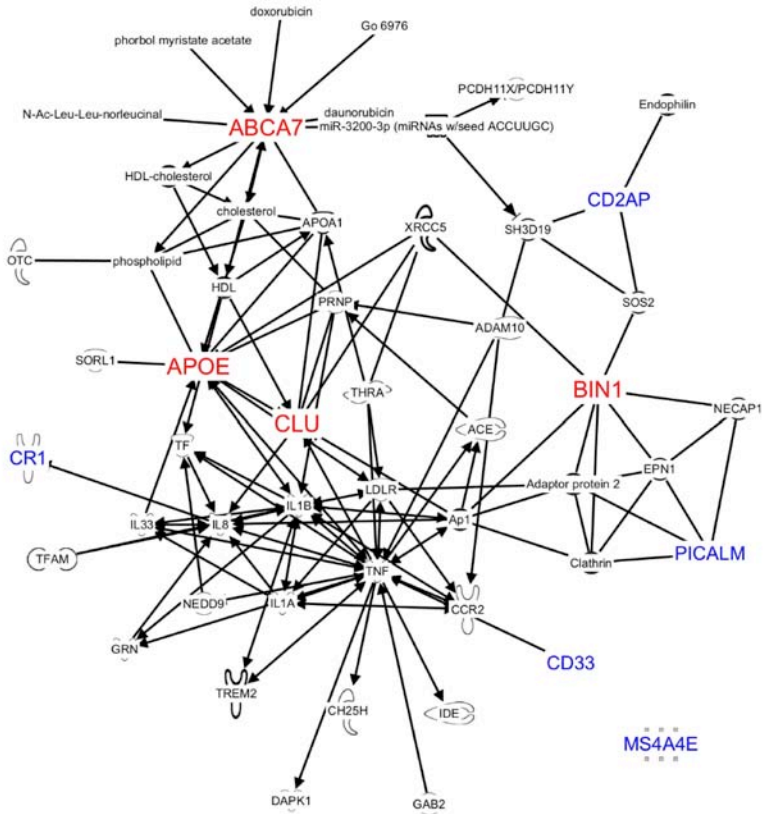

Figure 4: Network analysis predicts APOE, ABCA7, CLU and BIN1 as central players in Alzheimer's disease. We have constructed a LOAD-relevant gene network by interrogating the top 48 genes for LOAD (based on Alzgene) using Ingenuity Pathway Analysis Knowledgebase. Genes in blue and red are the top ten genes that we plan to interrogate in this proposal. (There is no known connection between MS4A4E and the genes in this network).

To perturb key parts of this network, we will generate mice carrying the human *APOE4* allele that are mutant for *Abca7*, *Clu* and *Bin1* (shown in red).

LOAD genes are expressed in multiple cell types in the cortex and hippocampus.

We have prioritized *APOE4*, *Abca7*, *Bin1* and *Clu* (the strongest candidate genes for LOAD) for early investigation in mice (Aim 3a). To establish further relevance for LOAD, we have used RNA in situ hybridization to determine the localization of *Abca7*, *Bin1* and *Clu* in control mice. All three genes have non-overlapping expression patterns in the cortex and the hippocampus. For example, in the CA1 region of the hippocampus, *Abca7* is expressed in both neurons and astrocytes (upper panels, Figure 5) whereas *Bin1* is expressed almost exclusively in neurons (lower panels).

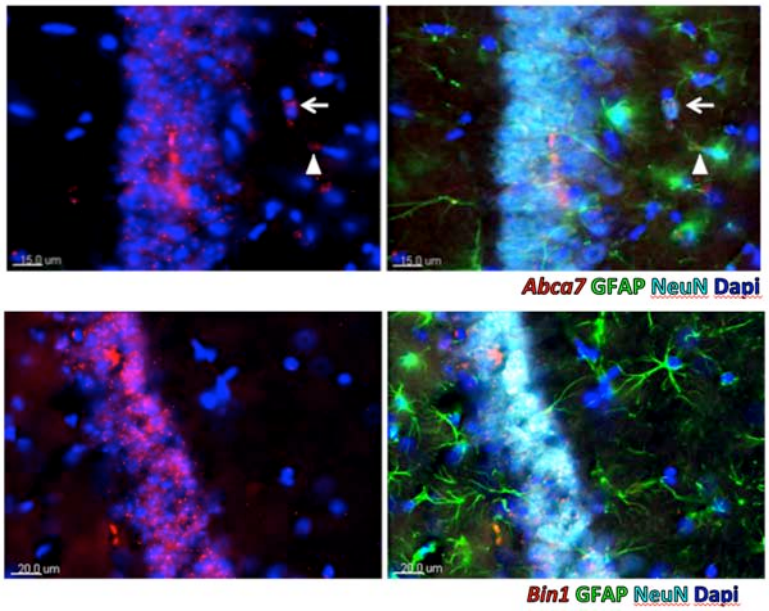

Figure 5. *Abca7* and *Bin1* show non-overlapping expression patterns in the CA1 region of the hippocampus. Representative images of localization studies for *Abca7* and *Bin1*. Riboprobes for *Abca7* or *Bin1* were cohybridized with antibodies for neurons (NeuN, light blue), astrocytes (GFAP, green) and nuclei (DAPI, dark blue). Arrows indicate neurons, arrowheads indicate astroctytes.

Genetic Interaction Modeling
We will carry out genetic interaction modeling of transcript data for specific combinations of mutations in mice. We have performed similar analysis in multiple studies using yeast as a model system, using both engineered gene knockouts [16], knockouts and multicopy perturbations [3], and mutations distributed in a recombinant population [14]. These studies used transcript data to predict patterns of differential expression and morphological phenotypes in novel mutant strains (Figure 6) and predictions were experimentally validated. More recently, we have applied the methods to signaling in flies (Carter *et. al*., under review) and are using the method to study the complex genetics of autoimmune disease in mice. The methods are designed for use in any model organism with engineered mutations and quantitative phenotype data, as described in this proposal.

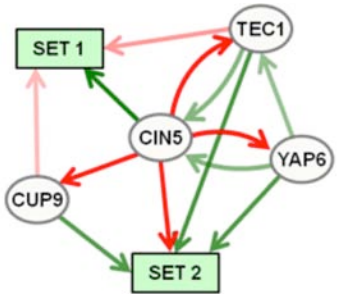

Figure 6. Network model of four yeast genes (ovals) interacting to affect two sets of co-expressed transcripts (boxes). Green (red) edges represent positive (negative) influences on gene activity or transcript levels. Adapted from Ref. [3].

Experimental Design and Methods

Specific Aim 1: To identify biological processes impacted by heterozygous and homozygous deletions of the ten strongest candidates genes for LOAD using RNAseq.

Rationale
Current predictions of the function of genes associated with LOAD are extremely difficult because a comprehensive picture of the pathways in which they function in regions relevant to AD is not available. To make accurate predictions as to the role of these genes in LOAD it is necessary to identify the pathways/networks in which they function. To identify these pathways, we will profile (using RNA-seq) AD-relevant brain regions of mice carrying mutations in each gene. Pathways/networks in which each gene functions will be altered compared to control mice. **By profiling the strongest genes associated with LOAD, we will identify pathways/networks likely critical for the development of AD.**

Experimental Cohorts
We will profile nine genes associated with LOAD in the first round of gene expression profiling (summarized in Table 1). At The Jackson Laboratory, we have available to us one of the largest collections of mutant mice in the world. In addition JAX is part of the Knockout Mouse Project 2 (KOMP2, see letter of support from Dr. Steve Murray). This provides us easy access to all the necessary mouse strains for this project (see Table 1).

All cohorts of mice will be on the C57BL/6J genetic background. This will minimize strain background effects in the gene expression profiling data. Alleles for some genes are on the C57BL/6N background (e.g. those obtained from KOMP2). For these mice, we will breed them to C57BL/6J (B6) for at least 5 generations (N6). We have started collecting and backcrossing those strains that are not on a B6 background. We anticipate having all necessary strains on the correct genetic background within the first year of the grant.

Profiling of LOAD genes will be performed on a sensitized strain background. This is because it is possible that a mutation in a single gene will not perturb all AD-relevant pathways. Therefore we are using a mouse strain that carries the human allele of *APOE4* and is likely sensitized to develop AD but shows no overt symptoms of AD such as plaque deposition or neuronal cell dysfunction. The *APOE4* allele confers the largest increase in risk of developing AD but alone is not sufficient to cause the disease. We will separately profile the sensitized strain to identify the “background” of networks/pathways perturbed by these mutations. A sensitized approach is a widely used strategy in mouse genetics and maximizes the likelihood of constructing AD-relevant pathways/networks. In addition, it is possible that a mutation in one LOAD gene on the sensitized background is sufficient to develop AD-relevant phenotypes and this would be of great benefit as a new AD model.

Table 2. Mutant strains to be profiled in the first round of experiments

Gene Symbol (Alzgene rank)	***Mutant Mouse strain***	***Genotypes for Profiling***	***Source***
APOE (1)	C57BL/6J.129P2-*APOE*$^{tm3(APOE*4)Mae}$	Hemi	Purchased from Taconic
BIN1 (2)	C57BL/6J.129S6-*Bin1*^{tm1Gcp}	Het only**	Kindly provided by Dr. Prendergast
CLU (3)	C57BL/6J.Cg-*Clu*^{tm1Jakh}	Het and Homo	Obtained from JAX mice and services
ABCA7 (4)	C57BL/6J.*Abca7*$^{tm1a(EUCOMM)Hmgu}$	Het and Homo	Available through KOMP2
PICALM (6)	C57BL/6NJ.*Picalm*$^{tm1e(KOMP)WTSI}$	Het and Homo*	Available through KOMP2
CD33 (8)	C57BL/6NJ.*Cd33*$^{tm1(KOMP)Vlcg/J}$	Het and Homo	Available through KOMP2
CD2AP (10)	C57BL/6J.129X1-*Cd2ap*^{tm1Shaw}	Het only**	Obtained from JAX mice and services
SORL1 (11)	C57BL/6NJ.*Sorl1*$^{Gt(pGT1tlr)1Ucd}$	Het and Homo	Available through KOMP2
TREM2 (-)	C57BL/6J.129P2-*Trem2*$^{tm1(KOMP)Vlcg}$	Het and Homo	Available through KOMP2

CR1 (ranked 5th) is not present in mice and so is not included at this point (see potential problems and improvements below). MS4A6A (ranked 7th) and MS4A4E (ranked 9th) are part of a complex gene family. It is not clear at this point which are the best transcripts to target and so are not included at this stage. Hemi=Hemizygous, Het = Heterozygous, Homo=Homozygous.
**A partial lethality has been reported for PICALM-/- mice but we anticipate being able to generate sufficient mice for profiling.*
**Mice that are homozygous null for either *Bin1* or *Cd2ap* die prior to 6 months of age and so will not be able to be profiled.

In total, two control cohorts and fourteen experimental cohorts will be generated and profiled in Aim 1. These include C57BL/6J (control), C57BL/6J.129P2-APOE$^{tm3(APOE*4)Mae}$ (sensitized strain) and eight cohorts that are either homozygous null or heterozygous for the LOAD gene on the sensitized genetic background (e.g. C57BL/6J.h*APOE4*$^{+/-}$.*Bin1*$^{+/-}$). To further minimize irrelevant gene expression differences, all mice will be (i) bred and aged in the same mouse room to be environment matched, (ii) female (sex matched), and (iii) harvested at the same age (6 months old). We have initially selected 6 months of age as this is when we anticipate early changes. Understanding this early stage of AD provides the largest window for therapeutic intervention. However, we will breed additional cohorts and age to one year in order to have the capability to profile these mice at a later date should it be necessary.

Gene expression profiling and identification of perturbed networks/pathways
For each of the sixteen cohorts to be profiled, we will assess both the cortex and the hippocampus in 6 biological replicates (192 samples in total). We will obtain tissue from one hemisphere, leaving the other hemisphere for validation of results and assessment of AD-relevant phenotypes (see below). It is not expected that the right and left hemispheres will show significant differences that could affect the interpretation of results. Tissues will be harvested and provided to The Jackson Laboratory's Genome Technologies service for RNA isolation, library preparation and sequencing. They perform these tasks routinely (see letter of support, Dr. Doug Hinerfeld, Associate Director). Samples will be prepared separately and then combined into pools of 6 for sequencing on the Illumina HiSeq.

We will use the open-source Galaxy software platform for the analysis of RNAseq data. Galaxy is a curated suite of tools for performing RNAseq analysis [17], and has been installed and is supported at The Jackson Laboratory. We will use the Groomer tool to assess the quality of the FASTQ-formatted data from the Illumina-supplied data pipelines. Groomer will parse the sequencer output and calculate summary statistics for base quality scores and nucleotide distributions across reads. Reads will be trimmed or discarded when nucleotide distributions become skewed, suggesting the depletion of nucleotide-specific reagents. Reads will be filtered by quality scores, with a low tolerance for mismatches due to our isogenic inbred strain backgrounds. The Tophat tool will be used to align RNA reads to the genome. Tophat is an implementation of the Bowtie alignment tool specifically designed for RNAseq and optimized for paired-end sequencing. The output, in sequence alignment map (SAM) format, will be assembled into transcripts using the Cufflinks tool. Cufflinks will match reads to putative transcripts and infer alternate splicing variants, and then quantify expression levels for each transcript [18,19]. Replicates will be individually analyzed for quality control and then averaged. However, improved tools are emerging all the time and we quickly incorporate the latest advances into our analysis. Throughout this process we will draw on the expertise of our collaborator, Dr. Gary Churchill, to address mouse-specific questions of alignment and quantification (see letter of support).

We will identify differentially expressed (DE) genes by comparing gene expression profiles from mice either heterozygous or homozygous mutant for a LOAD gene to profiles from both the control and sensitized strain. To identify processes perturbed by each mutation, we will use a variety of public and private resources including Database for Annotation, Visualization and Integrated Discovery (DAVID [20]), Gene Set Enrichment Analysis (GSEA [21]) and Ingenuity Pathway Analysis (IPA, www.ingenuity.com). We have all the necessary skills and experience to perform these analyses ([22,23] and letter of support, Dr. Simon John). Key findings will be confirmed using a combination of RNA in situ hybridization and immunofluorescence performed on sections from the remaining brain hemisphere. Members of networks/pathways will be localized to specific cell types using antibodies/probes specific to neurons (e.g. NEUN, *THY1*), astrocytes (GFAP, VIM), microglia/monocytes (*CD11B*, *AIF1*) and vascular endothelial cells (TEK). We will also assess Aβ deposition using antibodies including 4G8, 6A10 and Thioflavin-S.

This type of data does not currently exist for these important genes and will be extremely valuable to us, and others, as we move forward. To maximize the utility of this data, it will be made available both in Gene Expression Omnibus (GEO) (NCBI) and *GeneWeaver*, a user-friendly, online database (see Innovation, [13]).

Potential outcomes and interpretation of results

Analyses of currently available functional data predict LOAD genes fall into a number of different processes relevant to AD including Aβ aggregation and clearance (e.g. *APOE*, *CLU*, *PICALM*), Aβ degradation (*CD33* and *EPHA1*), lipid metabolism (e.g. *ABCA7*), and immune responses (*TREM2* and *CD33*). We will be able to test these predictions and, additionally, our unbiased approach will identify a comprehensive set of process/networks in which these genes function. For at least some genes (if not all) we expect their role to be far more detailed and complex than is currently believed.

Six biological replicates for each cohort will provide the sensitivity to detect even small gene expression changes between test and control cohorts. By profiling both the cortex and hippocampus from the same brains we will be able to identify region-specific functions. Importantly, our validation will allow us to determine cell-type specific functions. The use of RNA-seq provides the ability to assess not just the overall expression of genes, but individual transcripts for each gene, and so we can determine whether mutations in LOAD genes differentially regulate specific genes at the exon level.

We have selected the strongest candidate genes implicated in LOAD. However, human studies alone are not always able to definitively associate a gene with a disease. For instance, genome-wide associations studies associate a probed variant with a disease. In some cases, this variant lies within a gene but in many cases is close to a gene or between genes. In this case, the closest gene is considered a good candidate for the disease. This type of analysis is rarely sufficient to identify the causative variant and follow up studies are required to identify putative variant alleles for further validation. Our experiments in mice will provide new and extensive data for the role of LOAD genes in AD. As well as identifying novel networks/pathways, we will also be able to identify how the perturbation of each LOAD gene affects the networks/pathways that have been previously implicated in LOAD. We may also detect instances in which a LOAD gene does not perturb any networks/pathways that have previously been implicated in AD, which may suggest novel avenues for research based on its observed effects.

Potential problems and improvements

We do not anticipate major problems with generating the cohorts of mice, harvesting and profiling the tissue. It is possible that on a sensitized background homozygous mutations in some genes will lead to unexpected lethality prior to 6 months of age. If this is the case, we will profile these genes using only the control background strain in the absence of the APOE4 sensitizer. Conversely, it may be that mutating a single gene on the sensitized background is not sufficient to perturb any pathways. In this case we will further sensitize the background using a mutation in the mouse *Psen1* gene (linked with early onset AD). *Psen1*$^{M146VKI/\ M146VKI}$ mice do not show overt symptoms but show increased susceptibility to kainate-induced degeneration of CA1 and CA3 hippocampal neurons [24]. In this scenario, we would separately profile this 2-gene sensitized strain to identify the “background” of networks/pathways perturbed by these mutations.

Our current proposal does not include analysis of *CR1*. This is predicted to be an important gene in LOAD [25] but assessing the function of this gene in mouse is complicated. Mice do not have a direct ortholog of *CR1*, and instead produce a protein from the *Cr2* gene that is expected to act in a similar manner to the human CR1 protein [26]. It is thought that different splice forms (long or short forms) of *CR1* affect susceptibility to AD [25]. In collaboration with Genetic Resource Sciences at The Jackson Laboratory, we are currently genetically engineering mice that encode long and short forms of the human *CR1* protein. Given the difficulties of generating these mice we are not including them in the current proposal. However, we anticipate that we will be able to incorporate these mice into our studies within the granting period.

New genes are being associated with LOAD regularly (for example, *Trem2* was recently identified [9-11,27,28]) and it is likely that additional strong candidates are identified during the granting period. Our collaborators and key personnel on this project include Prof. Julie Williams who has been at the forefront of the identification of genes associated with LOAD. Her work has directly contributed to the identification of at least 10 LOAD genes. Wherever possible, we will incorporate these new genes into our pipeline. The resources we have available to us at The Jackson Laboratory make us well placed to quickly react to new discoveries.

Specific Aim 2: To map complex genetic pathways relevant to LOAD with computational modeling of double-knockouts.

Rationale
We will use data generated in Aim 1 to classify the candidate LOAD genes into subgroups that affect functional gene expression modules in specific tissues. We propose to use genetic interaction analysis to validate that the grouped genes co-function and model how they co-function. Genetic interaction, or epistasis, analysis is a method of systematically mapping the regulatory influences from gene to gene and from gene to phenotype. It is a classical method to order genes in pathways [29] and identify cases of genetic buffering [30]. By analyzing pair-wise combinations of our gene knockouts, we expect to understand if and how these LOAD genes form a network that regulates downstream transcript expression and, in turn, the biology underlying LOAD. The result will be polygenic models that can be validated using functional testing, data that will also be incorporated into further iterations of the modeling to improve their accuracy.

Approach
We will follow a two-step approach in mapping the interactions between LOAD genes. First, we will identify likely interacting genes. Next, we will construct targeted double-mutant strains and collect transcript data. We will then use computational genetic interaction analysis (CAPE) to model how the LOAD genes interact to affect the biology of LOAD.

Identification of Co-Functional LOAD Genes
We will first identify LOAD mutants and their effect on biological processes for genetic interaction analysis. To identify sets of transcripts affected by the same LOAD mutants, we will use singular value decomposition (SVD) to cluster groups of co-expressed transcripts and associate these groups with subsets of the samples (Figure 7) [31,32]. These gene groups will be assessed for common functional annotations using DAVID and other methods cited in Aim 1. Since each sample is a LOAD mutant with a tissue type, this analysis will allow us to identify subgroups of LOAD genes that affect specific biological processes in the cortex and/or hippocampus (for examples see Aim 1).

SVD will be used since it is an unsupervised method that: allows the inference of overlapping transcript groups; is scalable to large data sets; associates groups of samples with groups of transcripts; is a strict matrix operation with stable results that requires no parameter tuning; finds both strong and faint signals in the data; and filters noise. We are especially interested in identifying transcript groups for which a subset of mutants causes common effects (e.g. *Picalm* and *Clu* in Figure 7). In these cases we will hypothesize that the perturbed genes co-regulate a common function, determined by enriched annotations of the affected transcripts.

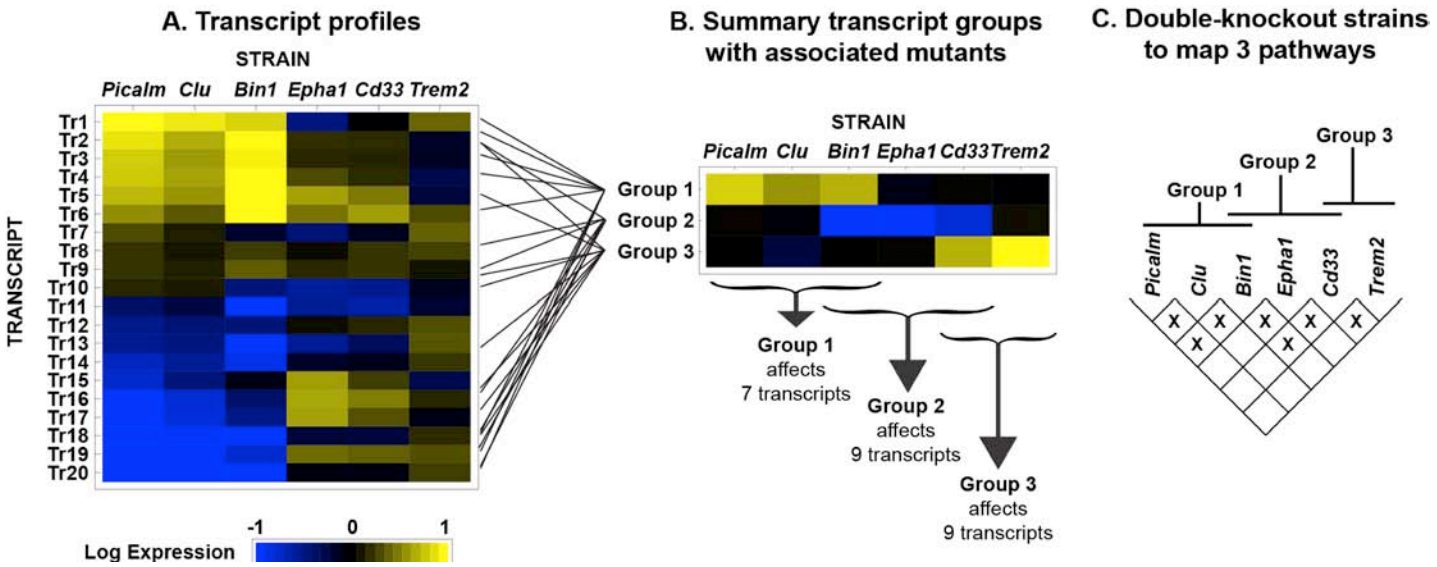

Figure 7. Clustering mutants by RNA expression, illustrated with artificial and simplified data. In reality the procedure will be applied to all 10 of our mutant strains and performed separately for both cortex and hippocampus samples. **(A)** Twenty hypothetical transcript profiles (Tr1-Tr20) across five mutant strains. **(B)** SVD analysis yields three gene groups with corresponding patterns of expression. Transcripts that significantly exhibit each pattern (either positively or negatively) are mapped by black lines to corresponding summary profiles. Transcripts in each group (200-300 are expected) will be queried for common functional annotations to identify affected biology. **(C)** Seven double-knockout strains (denoted X) to be made by combining all knockouts in each Group. For example, all three double-knockouts of *Picalm*, *Clu,* and *Bin1* mutants will be made to assess their combinatorial effects on transcripts in Group 1.

Genetic Interaction Analysis to Construct Network Models of LOAD

To determine the genetic architecture of the polygenic co-regulation of a transcript group, we will construct targeted pair-wise mutant strains and perform RNA-seq analysis identical to Aim 1. We will then be able to compare the single and double mutant effects to construct genetic interaction networks regulating LOAD biology. Six female double-mutant mice will be profiled at 6 months of age using the same protocol as in Aim 1. Although we do not know the precise number of promising mutant combinations we will identify, we expect to prioritize approximately 20 double-mutant samples in total, including both cortex and hippocampus tissues. The data will be combined with the single-mutant transcript data from Aim 1 for a systematic analysis of how the LOAD genes interact to affect groups of transcripts.

Our approach will be based on CAPE that casts genetic interactions in terms of gene-to-gene influences that correspond to activation or repression of one gene by another. The genes, in turn, influence downstream phenotypes such as transcript expression and other quantitative phenotypes. Thus the model quantifies both direct gene-to-phenotype effects and indirect effects manifest as genetic interactions such as buffering, suppression, and synthetic effects. The use of multiple phenotypes (patterns of transcript expression) related to LOAD greatly constrains the solutions and thereby resolves ambiguities in causal interpretations of interactions that often arise when single phenotypes are analyzed in isolation. In contrast to Bayesian inference approaches, this method allows the inference of cyclic graphs that correspond to positive and negative feedback loops (Figure 6). Feedback can be crucial for accurate prediction of the effects of system perturbations that modify the activity of genes [3] and should therefore improve selection of potential therapeutics. The strategy is outlined in Figure 8 and mathematical details are sketched in Box 1. To date, we have used this method to successfully analyze pathways in yeast model systems (Preliminary Data) [3,14,16]. With the proliferation of mouse knockout mutants and their availability at The Jackson Laboratory, we are now applying this strategy to mouse models of disease.

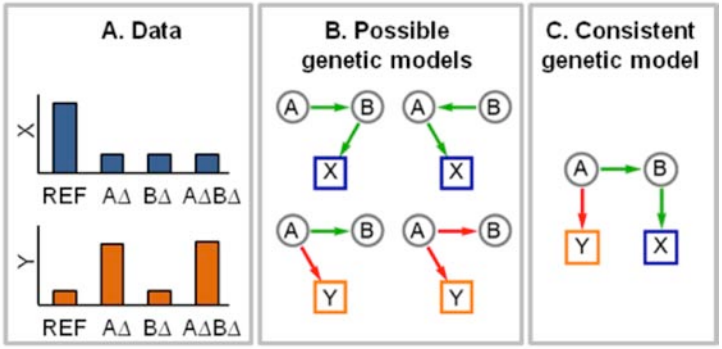

Figure 8. Use of pleiotropy to interpret genetic interactions. In our case the reference (REF) strain will be the *B6.APOE4* sensitized strain. **(A)** Effects of two knockouts, A and B, in isolation and combination on two transcripts, X and Y. **(B)** Possible model interpretations in terms of gene activities for A and B on transcripts X and Y, respectively. **(C)** The simplest model of influences between A and B that is consistent with X and Y.

Potential outcomes and interpretation of results

The expected outcomes of our analysis are models of how candidate genes interact to affect different aspects of the biology that underlies LOAD. We expect to observe many additive effects, corresponding to LOAD genes operating in distinct pathways to contribute to disease. We also expect to find multiple genetic interactions between subsets of genes, forming cliques or near-cliques of densely interacting nodes in the network models. We expect these models to have cortical or hippocampal specificity, identifying dysregulation of tissue-specific processes that contribute to LOAD. These network subsets will represent genes involved in the same pathway. By analyzing the direction of activation and suppression for each gene pair, we will infer pathway ordering based on standard rules [29].

Potential problems and improvements

Our modeling approach primarily relies on transcript data as a quantitative phenotype that represents LOAD biology. While we expect to see many different patterns of differential expression due to our genetic perturbations, these patterns might not map to known biological functions. In this case, we will use a more supervised analysis of the transcript data by identifying groups of transcripts that correlate with our physiological measures. Transcript levels within a group will be averaged to replace the SVD-derived patterns we proposed and modeling will continue as planned. Another possible outcome is that we do not observe any genetic interactions at the transcript level. While this would render some aspects of our analysis superfluous, the models we derive will still be valid for constructing mouse models of LOAD in Aim 3. Furthermore, simple interaction patterns might be more appropriately modeled by simple logical models [33]. Alternatively, the data might encompass too much biological complexity for our network analysis. This would be manifested as the inability to find a consistent model to explain the data. In this case, we select sets of transcripts from that data that are most relevant to the physiological phenotypes (e.g. transcripts correlated with Aβ aggregation and clearance in the hippocampus) and perform a more focused model that addresses specific biological processes in one tissue type.

Box 1: Mathematical modeling. We will use a modified regression approach that was developed for multivariate data such as transcript levels [3,16] or other quantitative phenotypes [14]. This comprises a matrix decomposition of the data matrix with knockout strains in rows and phenotypes (X, Y, Z) in columns:

$$\begin{pmatrix} X^{REF} & Y^{REF} & Z^{REF} \\ X^{A\Delta} & Y^{A\Delta} & Z^{A\Delta} \\ X^{B\Delta} & Y^{B\Delta} & Z^{B\Delta} \\ X^{A\Delta B\Delta} & Y^{A\Delta B\Delta} & Z^{A\Delta B\Delta} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & g_B^{A\Delta} \\ 1 & g_A^{B\Delta} & 0 \\ 1 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} x_0 & y_0 & z_0 \\ x_A & y_A & z_A \\ x_B & y_B & z_B \end{pmatrix} + \varepsilon$$

This is a linear model fit of how each gene influences each phenotype (x_A, x_B, etc), with variables added for modifications of each gene's activity when other genes are knocked out (e.g. $g_B^{A\Delta}$ is Gene B's activity when Gene A is deleted). This procedure straightforwardly extends to more than two knockout genes when data is available for all pair-wise combinations of knockouts [14]. The modified activity variables are then used to compute gene-to-gene influence coefficients that model linear influences of each gene on the activity of all others, by fitting a model of the form $g_A = g_{A0} + m_{AB} g_B$ for all knockout genes across all strains [3,16]. This formalizes the idea that, for example, a reduction in Gene A's activity in a Gene B knockout strain corresponds to a *positive* influence from Gene B to Gene A. Bootstrap cross-validation will be used to estimate the statistical significance of each edge [16] and we will systematically compare our interacting models to additive models without interaction [3].

Specific Aim 3: To generate mouse models for LOAD and perform early validation of novel therapeutic targets.

Rationale

No treatments for the prevention of LOAD exist. The identification of targets for potential treatments is severely hampered by the lack of understanding of how variations in LOAD genes interact to cause LOAD. Our preliminary data has identified combinations of mutations that are predicted to interact to cause LOAD. Our iterative computational modeling in Aim 2 will also generate networks of interacting genes that can be functionally tested (see Figure 1). The work in this aim will be two fold. First, we will functionally test the genetic models by determining whether mutations in combinations of LOAD genes are sufficient to cause AD-like phenotypes in mice. Second, our models can identify those genes/pathways that would be predicted to have a protective effect on AD-like phenotypes. These novel targets do not need to directly interact with LOAD genes but may act in an indirect manner and so would not necessarily be identified by other methods. **Ultimately, we aim to generate the first models for LOAD and identify genes that when modified can reduce/protect from AD-relevant phenotypes.**

Experimental cohorts

Aim 3a: Based on pathway/network analysis we hypothesize that deleterious variations of *CLU*, *ABCA7* and *BIN1* in the presence of APOE4 will be sufficient to cause AD-relevant phenotypes in mice (see Preliminary data, Figure 5). To test this, we will generate two cohorts of mice. In the first cohort, we will generate 10 mice (5 males and 5 females) homozygous for mutations in *Clu* (chr18) and *Abca7 (chr10)* and a heterozygous mutation in *Bin1* (chr18). All mice will also be hemizygous for the human *APOE4* allele.

Aim 3b: Iterative computational modeling (Aim 2) will generate high quality gene networks for functional testing. We will generate similar cohorts of mice to those described in Aim 3a above to test these networks. We will test as many models as is possible within the timeframe of the grant but anticipate being able to test at least three networks.

Assessment of Alzheimer's disease-relevant phenotypes

Mice will be aged to 12 months and assessed for initially behavioral phenotypes before being harvested and processed for histological examination and molecular profiling:

Behavioral testing: We will use a series of learning and memory tests that assess both short term and long-term memory. These will include the object recognition, social transmission, 'Y' and 'T' mazes and Morris water maze. We will draw on the expertise of our co-investigator, Dr. Elissa Chesler, for this work.

Histological examination: Multiple phenotypes will be assessed on the same brain tissue (e.g. plaque deposition by ELISA on the left hemisphere of the brain and immunofluorescence on the right hemisphere). Plaque burden will be assessed using a combination of ELISA and immunofluorescence (see Preliminary data). We will also assess neuronal cell loss, neurofibrillary tangles and glial activation. We have developed working conditions for antibodies that recognize neurons (e.g. THY1, NeuN), neurofibrillary tangles (phosphorylated TAU), astrocytes (e.g. GFAP), microglia (e.g. IBA1 and CD11b), and endothelial cells (e.g. TEK).

Molecular profiling: For each cohort, one hemisphere from three female mice will be selected for gene expression profiling of the cortex and hippocampus using the same methods described in Aim 1. This data will allow us to generate an accurate picture of the biological processes that are perturbed by mutations in the each of gene combinations. In addition, given the data was generated under exactly the same conditions it can be incorporated into our computational modeling in Aim 2 and allow further refinement of the gene networks.

Potential outcomes and interpretation of results

A key feature of AD we are trying to recapitulate in mice is the significant dysfunction and ultimately demise of neuronal populations in the cortex and hippocampus. This would be expected to lead to severe learning and memory deficits. This is lacking from current models and hampers our ability to test new drug targets. In this aim, we are testing a predictive model of LOAD based on current gene interaction knowledge (Aim 3a). We would expect at least some AD-relevant phenotypes to be present in these mice particularly in combination with the homozygous *Psen1* mutation. We will also test additional combinations of genes predicted through our

computational modeling (Aim 3b). We do not expect only one combination of gene mutations to cause LOAD. As is likely in humans, variations in different sets of genes are likely to cause similar outcomes.

Given the complexity of AD, we anticipate this being an iterative process and we may need to go through two or more rounds of modeling/validation. Our experiments are designed in such way that at each step, we will generate very valuable data that will contribute to our understanding of how LOAD genes functions. As many iterations will be completed within the time frame of this grant.

Potential problems and improvements

We are confident of our breeding strategy for generating cohorts of mice for Aim 3a. However, if no AD-relevant phenotypes are observed, we will generate a second cohort incorporating an additional sensitizer mutation, the homozygous mutation for *Psen1* (chr12, [24]). The use of this second sensitizer maximizes the likelihood of developing AD-relevant phenotypes.

It is possible that combinations of homozygous mutations cause premature lethality. Should this be the case, we will generate mice carrying heterozygous mutations. Heterozygous mutations cause functional deficiency (or haploinsufficiency) in many cases and may actually recapitulate the human condition. Recent data from the EuroPhenome project that shows that greater than 50% of genes show a phenotype when one copy of the gene is ablated (unpublished data, KOMP2).

In this proposal, we have chosen the uniformity of assessing heterozygous and homozygous knockouts. As we learn more, we will remain flexible to incorporate different types of mutations into our pipeline in the future. Firstly, it will likely be necessary to assess overexpression of some genes through transgenics. We are already using the transgenic form of APOE4 and as stated earlier we are separately generating transgenic mice carrying human CR1 isoforms. If other relevant transgenic mice become available we will consider incorporating them. We also have the capability at The Jackson Laboratory to generate additional transgenic mice. Secondly, we have the potential to assess more subtle variations such as amino acid substitutions (hypomorphs). Using a random ENU mutagenesis strategy we have identified point mutations in *Clu*, *Abca7*, *Bin1*, *Cd33* and *Epha1* and this resource continues to grow (Howell, John, unpublished). At this stage we do not believe enough is known about the role of LOAD genes to use this resource yet but as we learn more, we anticipate being able to incorporate these mutations into our study in the future.

Conclusion and Timeline

We propose ambitious but timely experiments to better understand the role of LOAD genes in Alzheimer's disease. The identification of the fundamental pathways perturbed in AD will provide the foundation to develop new drugs for human AD. We have put together experts in their field to maximize our chances of success. We are well placed to perform this work at The Jackson Laboratory given all the mouse resources at our disposal. However, we will not work in isolation - we have established important collaborations with human geneticists. Combining the strengths of human genetics/genomics with mouse genetics/genomics is key to developing preventative treatments for Alzheimer's disease.

Timeline

		Year				
		1	2	3	4	5
Specific Aim 1	**Gene expression profiling of LOAD genes**					
	Unify genetic background	■				
	Breed and age cohorts	■	■			
	Harvest tissue and perform RNA-seq		■			
	Analysis/database development		■	■	■	■
Specific Aim 2	**Iterative computational modeling**					
	Identify likely interacting genes		■	■		
	Breed and age double mutant strains		■	■	■	
	Generating models of LOAD			■	■	■
Specific Aim 3	**Early validation of new drug targets**					
	3A - testing of *Clu*, *Bin1* and *Abca7*	■	■	■		
	3B - testing of new networks			■	■	■
	Validation of possible drug targets				■	■

VERTEBRATE ANIMALS

1. In this proposal we will use mice to generate resources for performing mechanistic studies of human diseases. We will primarily use C57BL/6J, a common inbred strain of mouse. Due to unification of strain background (an important component of this proposal), other inbred strains, such as C57BL/6NJ are likely to be used. We will use the minimum number of mice possible and do not expect this to exceed 1200 mice per year (400 males and 800 females). Three major criteria impact the number of mice to be used. (i) We are collecting and generating mouse strains that will be will useful to us, and others, and will need to maintain these strains throughout the granting period. (ii) For accuracy in transcriptional profiling, it is essential that the genetic background of mice is uniform. Therefore we anticipate backcrossing 5-10 alleles onto C57BL/6J. To ensure experiments are completed in the time frame, backcrossing has begun in advance of the granting period. However, it's anticipated that backcrossing will continue in the first year of the proposal. (ii) The proposal involves complex breeding strategies to generate mice carrying combinations of breeding. Building cohorts of 6-10 females of the correct genotype combination is expected to involve approximately 50 mice per cohort. Cohorts of mice will need to be generated for single gene assessment (on the sensitized background, Aim1), pairwise comparisons (Aim 2) and multiple combinations (Aim 3).

2. Due to limitations of working with humans, mice are an ideal model system to use. The work described in this proposal can only be carried out in mice. As an animal model, they are at the forefront of gene-targeting technologies. In particular, the tools developed for genetic manipulation of ES cells, and the fully characterized inbred strains of mouse are a necessary part of this work.

3. The Jackson Laboratory is accredited by the Association for the Assessment and Accreditation of Laboratory Animal Care (AAALAC). Animals used in the project proposed in this application will be housed in the Research Animal Facility (RAF). The Laboratory Animal Health Services (LAHS) program headed by Linda Waterman, DVM includes four veterinarians and two veterinary pathologists. The *Veterinary Service* includes methods of prevention, control, diagnosis and treatment of diseases and injuries, as well as a routine health-monitoring program, through which mice are monitored four times a year for an extensive battery of microorganisms. The *Diagnostic Service*, supervised by James Fahey, DVM, PhD, provides diagnostic support for animal health services and research programs. The *Clinical Medicine* group headed by Bonnie Lyons, DVM evaluates sick animals, prescribe treatments, and follow up to assure that the treatments are effective. Clinical investigations are initiated as needed to follow up on potential disease outbreaks. Veterinary staff are available on-site or on-call after hours to respond to critical situations 24 hours a day.

Four to five adult mice (depending upon body weight) or a single breeding pair with preweaned young are maintained in polycarbonate cages (50 square inches floor area). Mice are fed *ad libitum* with the NIH-31 6% or 4% fat mouse diet. Cages are washed once a week and filled with an appropriate amount of sterilized white pine shavings. Water bottles are also washed weekly and filled with acidified (pH 2.5-3.0) water. Cage lids are covered with a non-woven polyester flat filter. Food hoppers, which are attached to the metal cage covers, are filled weekly. Each cage is checked daily for proper water delivery, food levels, and overall well-being of the animals. All animals receive clean housing on a weekly or bi-weekly basis depending on the type of racking system. The light cycle is 12 hrs light:12 hrs dark.

4. All procedures are reviewed and approved by our Institutional Animal Care and Use Committee. Mice are anesthetized prior to any invasive procedure (such as perfusion fixation) with a ketamine/xylazine mixture (99mg/kg ketamine; 9mg/kg xylazine, 0.06ml/10g body weight) by intraperitoneal injection. During recovery, mice are placed in a warm environment and checked until they have recovered from anesthesia sufficiently to right themselves and move about the cage. Then they are returned to the animal room. The mice are examined periodically during the next few days to assure that incisions are healing properly and that normal hydration is being maintained. A topical anesthetic is administered prior to retro-orbital bleeding.

5. Mice will be euthanized according to protocols consistent with the Panel on Euthanasia of the American Veterinary Medical Association. Embryonic and neonatal mice will be euthanized by decapitation. Older mice will be euthanized by carbon dioxide asphyxiation.

MULTIPLE PI LEADERSHIP PLAN

The PIs (Drs. Howell and Carter) are investigators at The Jackson Laboratory and provide highly complementary expertise for this project. Dr. Howell has been applying genetic and genomic strategies to study age-related neurodegenerative diseases including Alzheimer's disease. Dr. Carter has extensive experience in gene expression analysis, the study of genetic interactions, and computational modeling of genetic systems. Dr. Howell's expertise in neurological phenotypes and mouse genetics/genomics paired with Dr. Carter's expertise in computational genetics will provide the necessary experience and leadership to carry out the proposed work.

For this application, Dr. Howell will guide overall direction for laboratory work, participate in the design and implementation tasks, oversee the animal experiments, phenotyping, and tissue collection, and write manuscripts and reports as required. Dr. Carter will provide primary direction for the sequence data analysis and computational modeling, contribute to data acquisition decisions, and write manuscripts and reports as required.

Drs. Carter and Howell will meet with project staff biweekly to evaluate progress and discuss any issues that have arisen in their work. These meetings can often clarify processes and identify areas where others need to be consulted. They will also meet to discuss planning aspects, review data, actions necessary based on the staff meetings, and evaluate any budget questions.

The PIs expect to reach common agreement on management issues by thoroughly discussing and carefully considering the pros and cons of specific actions. Their backgrounds are highly complementary, and they do not foresee any disagreements that would negatively affect the proposed research. However, should a difference arise, Dr. Howell will have final say on decisions related to the laboratory experiments, and Dr. Carter will have final say on decisions related to RNA sequencing and the analysis of the sequencing data.

REFERENCES CITED

1. Ballard C, Gauthier S, Corbett A, Brayne C, Aarsland D, Jones E. Alzheimer's disease. Lancet 2011;377:1019-31.
2. Zhang YW, Thompson R, Zhang H, Xu H. APP processing in Alzheimer's disease. Mol Brain 2011;4:3.
3. Carter GW, Hays M, Li S, Galitski T. Predicting the effects of copy-number variation in double and triple mutant combinations. Pac Symp Biocomput 2012:19-30.
4. Jones L, Holmans PA, Hamshere ML, et al. Genetic evidence implicates the immune system and cholesterol metabolism in the aetiology of Alzheimer's disease. PLoS One 2010;5:e13950.
5. Seshadri S, Fitzpatrick AL, Ikram MA, et al. Genome-wide analysis of genetic loci associated with Alzheimer disease. JAMA 2010;303:1832-40.
6. Harold D, Abraham R, Hollingworth P, et al. Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. Nat Genet 2009;41:1088-93.
7. Okuizumi K, Onodera O, Tanaka H, et al. ApoE-epsilon 4 and early-onset Alzheimer's. Nat Genet 1994;7:10-1.
8. Liu L, Forsell C, Lilius L, Axelman K, Corder EH, Lannfelt L. Allelic association but only weak evidence for linkage to the apolipoprotein E locus in late-onset Swedish Alzheimer families. Am J Med Genet 1996;67:306-11.
9. Neumann H, Daly MJ. Variant TREM2 as Risk Factor for Alzheimer's Disease. N Engl J Med 2012.
10. Guerreiro R, Wojtas A, Bras J, et al. TREM2 Variants in Alzheimer's Disease. N Engl J Med 2012.
11. Jonsson T, Stefansson H, Ph DS, et al. Variant of TREM2 Associated with the Risk of Alzheimer's Disease. N Engl J Med 2012.
12. Bertram L, McQueen MB, Mullin K, Blacker D, Tanzi RE. Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. Nature genetics 2007;39:17-23.
13. Baker EJ, Jay JJ, Bubier JA, Langston MA, Chesler EJ. GeneWeaver: a web-based system for integrative functional genomics. Nucleic Acids Res 2012;40:D1067-76.
14. Carter GW, Hays M, Sherman A, Galitski T. Use of pleiotropy to model genetic interactions in a population. PLoS Genet 2012;8:e1003010.
15. Jankowsky JL, Fadale DJ, Anderson J, et al. Mutant presenilins specifically elevate the levels of the 42 residue beta-amyloid peptide in vivo: evidence for augmentation of a 42-specific gamma secretase. Hum Mol Genet 2004;13:159-70.
16. Carter GW, Prinz S, Neou C, et al. Prediction of phenotype and gene expression for combinations of mutations. Mol Syst Biol 2007;3:96.
17. Blankenberg D, Von Kuster G, Coraor N, et al. Galaxy: a web-based genome analysis tool for experimentalists. Curr Protoc Mol Biol 2010;Chapter 19:Unit 19 0 1-21.
18. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 2008;5:621-8.
19. Jiang H, Wong WH. Statistical inferences for isoform expression in RNA-Seq. Bioinformatics 2009;25:1026-32.
20. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009;4:44-57.
21. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences of the United States of America 2005;102:15545-50.
22. Howell GR, Macalinao DG, Sousa GS, et al. Molecular clustering identifies complement and endothelin induction as early events in a mouse model of glaucoma. Journal of Clinical Investigation 2011;121:1429-44.
23. Howell GR, Soto I, Zhu X, et al. Radiation treatment inhibits monocyte entry into the optic nerve head and prevents neuronal damage in a mouse model of glaucoma. J Clin Invest 2012;122:1246-61.
24. Guo Q, Fu W, Sopher BL, et al. Increased vulnerability of hippocampal neurons to excitotoxic necrosis in presenilin-1 mutant knock-in mice. Nat Med 1999;5:101-6.
25. Crehan H, Holton P, Wray S, Pocock J, Guerreiro R, Hardy J. Complement receptor 1 (CR1) and Alzheimer's disease. Immunobiology 2012;217:244-50.
26. Jacobson AC, Weis JH. Comparative functional evolution of human and mouse CR1 and CR2. J Immunol 2008;181:2953-9.

27. Jones B. Alzheimer disease: TREM2 linked to late-onset AD. Nature reviews Neurology 2012;9:5.
28. Niemitz E. TREM2 and Alzheimer's disease. Nature genetics 2012;45:11.
29. Avery L, Wasserman S. Ordering gene function: the interpretation of epistasis in regulatory hierarchies. Trends Genet 1992;8:312-6.
30. Hartman JLt, Garvik B, Hartwell L. Principles for the buffering of genetic variation. Science 2001;291:1001-4.
31. Alter O, Brown PO, Botstein D. Singular value decomposition for genome-wide expression data processing and modeling. Proceedings of the National Academy of Sciences of the United States of America 2000;97:10101-6.
32. Carter GW, Rupp S, Fink GR, Galitski T. Disentangling information flow in the Ras-cAMP signaling network. Genome Res 2006;16:520-6.
33. Liang J, Han J. Stochastic Boolean networks: An efficient approach to modeling gene regulatory networks. BMC Syst Biol 2012;6:113.