



OPEN

NREM delta power and AD-relevant tauopathy are associated with shared cortical gene networks

Joseph R. Scarpa¹, Peng Jiang², Vance D. Gao², Martha H. Vitaterna², Fred W. Turek² & Andrew Kasarskis^{3,4}✉

Reduced NREM sleep in humans is associated with AD neuropathology. Recent work has demonstrated a reduction in NREM sleep in preclinical AD, pointing to its potential utility as an early marker of dementia. We test the hypothesis that reduced NREM delta power and increased tauopathy are associated with shared underlying cortical molecular networks in preclinical AD. We integrate multi-omics data from two extensive public resources, a human Alzheimer's disease cohort from the Mount Sinai Brain Bank (N = 125) reflecting AD progression and a (C57BL/6J × 129S1/SvImJ) F2 mouse population in which NREM delta power was measured (N = 98). Two cortical gene networks, including a CLOCK-dependent circadian network, are associated with NREM delta power and AD tauopathy progression. These networks were validated in independent mouse and human cohorts. Identifying gene networks related to preclinical AD elucidate possible mechanisms associated with the early disease phase and potential targets to alter the disease course.

Alzheimer's disease (AD) emerges over the course of decades, involving a dynamic integration of genetic, environmental, and behavioral risk. A growing body of literature has elucidated the neurobiological changes in the preclinical phase of AD, noting significant neuropathologic and behavioral changes before the onset of cognitive decline. Studies of patients with dominantly inherited Alzheimer's disease have demonstrated that amyloid-beta and tau accumulate for many years before pathologic changes in memory and executive function^{1,2}, and A β levels peak around the time clinical symptoms manifest³. These early stages of AD also are accompanied by widespread multi-region loss in synaptic density and neuronal populations³. This neuropathologic evidence complements the data demonstrating a complex neuropsychiatric symptomatology in the early stages of AD⁴⁻⁷. Recent studies have shown that the preclinical phase includes anxiety-depressive symptoms⁸ and increased loneliness^{9,10}. Other work has noted significant changes to sleep-wake cycle and its influence on AD pathology¹¹⁻²¹. This body of literature is highly concordant with those of Parkinson's disease and Huntington's disease that demonstrate pathologic and behavioral changes during the prodrome²²⁻²⁸. These characteristics may indicate a generalizable feature of neurodegenerative disease and may point to opportunities for early intervention.

Sleep and circadian disruption is a significant cause of morbidity in AD. This stressor is transmitted through the patient's social and familial network, affecting the mental health of families and caretakers²⁹⁻³². Often sleep and circadian disruption leads to the institutionalization of patients in long-term care facilities that more reliably offer 24-h care. Though sleep disruption plays a significant role in late-stage AD, the relationship between sleep and dementia is highly intertwined throughout the progression of Alzheimer's disease^{16,18-20,33}. Epidemiologic evidence indicates that sleep deprivation may increase dementia risk, and dementia often leads to sleep deprivation, potentially creating a feedback loop^{13-16,18,20,33,34}. A number of sleep changes have been associated with AD pathogenesis^{11,14,15,35-37}. Recent work has demonstrated a reduction in NREM sleep in preclinical AD, pointing to its potential utility as an early marker of dementia³⁵. Reduced NREM in humans is associated with increased atrophy and A β accumulation in the medial prefrontal cortex and tau accumulation in specific Brodmann areas in the frontal, temporal, and parietal brain regions^{35,38-40}. This work shows particular sleep features that manifest in the preclinical phase and are associated with specific elements of AD pathology, but the molecular mechanisms underlying these neurophysiologic markers are poorly understood. Understanding the biological basis of these

¹Department of Anesthesiology, Weill Cornell Medicine, New York, NY 10065, USA. ²Center for Sleep and Circadian Biology, Department of Neurobiology, Northwestern University, Evanston, IL 60208, USA. ³Icahn Institute for Genomics and Multiscale Biology, Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. ⁴Department of Population Health Science and Policy, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. ✉email: Andrew.kasarskis@mssm.edu

complex features of preclinical AD may reveal insights into pathogenesis of preclinical Alzheimer's and novel opportunities for therapy.

The motivation of the present work is to corroborate the relationship between NREM slow-wave activity (SWA) and cortical tau pathology at the molecular level and characterize the shared molecular basis of these intertwined phenotypes³⁵. We use data from two extensive public resources, a human Alzheimer's disease cohort from the Mount Sinai Brain Bank (N = 125) reflecting AD progression⁴¹ and a (C57BL/6J × 129S1/SvImJ) F2 mouse population in which NREM delta power was measured (N = 98)⁴², to test the hypothesis that specific gene networks in the prefrontal cortex are associated with both NREM sleep and AD-related tau pathology. These large-scale studies catalogue extensive genetic and molecular information across species—rich public resources that can be used to investigate specific biological questions not considered in the original studies. By studying genetic, environmental, and phenotype data across species, we can characterize how genetic variation and gene expression integrate information in gene networks and identify molecular pathways robustly associated with both NREM sleep and tau development.

Methods

Multiscale data generation in mouse and human cohorts. We previously described in full detail the experimental design, quality control, and measurements of genetic, gene expression, and phenotypes in ninety eight (98) members of (C57BL/6J × 129S1/SvImJ) F2 mouse population⁴². We used these mouse data for our analyses in the present study, and all animal experiments that generated these data were approved in advance by the Institutional Animal Care and Use Committee at Northwestern University and were in compliance with the Federal Animal Welfare guideline. Briefly, two hundred and eighty three (283) behavioral, affective, and electroencephalographic traits were measured, including NREM delta power. Mice starting at 7 weeks of age were subjected to a battery of behavioral test and at 10 weeks of age were surgically implanted with electroencephalogram (EEG) and electromyogram (EMG) electrodes for sleep/wake recordings. Two stainless steel screws (Small Parts, Miami Lakes, FL) serving as the EEG leads were screwed into the skull and placed in the cerebral cortex with one screw located 1 mm anterior to bregma and 2 mm lateral to the central suture (somatomotor areas), and the other (i.e., the reference electrode) at 1 mm anterior to lambda and 2.5 mm lateral to the central suture (visual areas). As is typical for studies in mice, one-channel EEG signals were used here as a proxy to describe the characteristics of cortical frequency spectrum in an individual animal. Two weeks after the surgery, three episode of continuous EEG/EMG recordings were made, including (1) a 24-h undisrupted baseline (ZT0 to ZT0; ZT0 = light on), (2) 6 h of sleep deprivation (starting at ZT2; i.e., 2 h after the end of baseline) followed by 16 h recovery period, and (3) sleep/wake after 1 h (ZT5–ZT6) of restraint stress given a week after the sleep deprivation. EEG/EMG recordings were manually scored in 10-s epochs into NREM, REM, and wake, and then quantified into sleep phenotypes. For our current analysis, we focused on the relative NREM delta power during the 24 h baseline sleep, the phenotype referred to as “BL.24h.nrem.delta” throughout the original manuscript. The power spectrum analysis was done epoch-by-epoch. Raw power of the delta band (1–4 Hz) was averaged for all NREM epochs in the 24-h baseline and was then expressed as a proportion of the epoch-averaged NREM total power of the 1–30 Hz range during the same period. The normalization to total power removes the variations in the EEG signal amplitude among individual animals, which are largely technical rather than biological variations. Our phenotype of interest, the relative NREM delta power during baseline reflects, the intensity of NREM delta activities, which is a commonly used marker of sleep homeostatic drive and is highly heritable^{43–45}. A closely related phenotype, NREM delta energy, which is a function of delta power and the amount of slow-wave sleep, is not analyzed here or in our previous study. Finally, as animal cohorts were found to have a widespread impact on the 283 phenotypes measured in this mouse population (ANOVA P values significantly deviate from a uniform distribution and are distorted toward smaller values), linear adjustments were used to remove cohort effects for all phenotypes. For each phenotype, phenotypic values of each animal were fitted in a linear regression model with cohort as the independent variable, and intercept + residuals were taken as the adjusted values. This linear procedure was taken because it is conceptually simple and compatible with subsequent analyses (e.g., correlations) which are also linear.

All mice were genotyped using the Affymetrix MegAllele genotyping mouse 5K SNP Panel, of which 2458 were polymorphic between the C57BL/6J and 129/SvImJ inbred strains. To measure gene expression, Affymetrix GeneChip Mouse Genome 430 2.0 Array was used. Mice were left undisturbed for two weeks before euthanasia and tissue collection. All dissections were performed between ZT6 and ZT7, and four brain regions were collected, including frontal cortex (cortical tissue rostral to the striatum), hippocampus, thalamus, and hypothalamus. Only the expression data in the cortical tissue were used the current study. Raw data were normalized using robust multiarray averaging with quantile normalization and consequently log₂-transformed. We adjusted data for effects of technical covariates, including RNA processing batch, array batch, and animal cohorts. Principal components analysis was used to capture the unknown source of variations, and data were further adjusted by treating principal components as covariates if they were not associated with any genomic loci [logarithm of the odds (LOD) < 3]. Lastly, multiple probesets mapping to a single gene were reduced to the gene level by calculating their median. The bottom 10% of genes by variance were removed to reduce noise in calculating gene coexpression, and samples with an interarray correlation two standard deviations away from the mean were considered outliers and removed. Gene expression and phenotype data are publicly available at GSE109112.

Data from the Alzheimer's disease cohort was collected by the Mount Sinai Brain Bank, and full description of the experimental methods and data generation procedures have been previously published⁴¹. These data include 1053 postmortem brain tissues from 125 human brains. Gene expression was measured and merged from two Affymetrix microarray platforms, Human Genome U133A and U133B. Probes measured on both platforms were averaged before the gene expression matrix was corrected for covariates using linear regression. For each member

of the cohort, clinical dementia rating (CDR), diagnostic certainty score through the Consortium to Establish a Registry for Alzheimer's disease (CERAD), Braak Score, and estimates of neuritic plaque (NPL) and neurofibrillary tangle (NFT) density were measured using a previously published protocol⁴⁶. This cohort includes individuals expressing the full spectrum of clinical and histopathologic phenotypes, from normal to severe, in order to correlate variations in gene network expression with clinical and histopathological status and disease progression. Gene expression data is publicly available at GSE84422, and all data can be accessed through Synapse.

Calculating coexpression and module enrichment to identify functional molecular relationships. Both previous studies which we primarily examine in this work use weighted gene coexpression network analysis (WGCNA) to calculate gene coexpression associations in high-dimensional microarray data^{41,42}. We further analyzed these coexpression networks to identify molecular relationships between tauopathy and non-REM slow wave activity. These coexpression networks were previously generated using a similar generalized methodology. For each cohort, the correlation matrix was calculated between all genes. This correlation matrix was then raised to a power, β , to calculate the adjacency matrix⁴⁷, which was quadratically transformed to the topological overlap matrix to capture higher-level gene–gene relationships, like nearest neighbor associations⁴⁸. Hierarchical clustering revealed gene coexpression modules, which were assigned arbitrary colors for identification. Each color represents a different module, which was catalogued in the original manuscript. The first principal component of each module (module eigengene) was correlated with relevant phenotypes to determine associations between modules and traits. In the original F2 mouse study, twenty-six coexpression modules were identified. Notably, gene expression in three modules (“blue”, “green”, and “skyblue”) were found to be altered both by sleep deprivation and major depression. Network images were generated using Cytoscape v3.7.2⁴⁹.

To identify gene networks associated with both AD-relevant tauopathy and non-REM slow wave activity, we compared module memberships between all previously calculated coexpression modules in the frontal cortex of the (C57BL/6J \times 129S1/SvImJ) F2 mouse population and BA10 of the human AD cohort. Fisher's exact test was used to calculate module overlap and a module was considered conserved when Bonferroni-corrected p -value < 0.05 and Odds Ratio > 2 . Fisher's exact test was also used to investigate whether the two conserved modules were overrepresented with genes differentially expressed in non-demented individuals with AD histopathology when compared to an age- and sex-matched cohort that was clinically and histopathologically normal. Nominal p -values were reported since only two modules were investigated. Functional characterization of modules were calculated using DAVID^{50,51}, and Benjamini–Hochberg p -values are reported. Furthermore, enrichment using ChIP-seq and CLOCK-knockdown data was performed with Enrichr and corrected p -values were reported⁵². The ChIP-seq data was catalogued in the ChEA database, derived from an experiment examining CLOCK-binding targets in 293T cells⁵³. The knockdown data is publicly available from GSE50588, generated in GM19238 cell lines after siRNA knockdown, and catalogued in Enrichr under “TF Perturbations Followed By Expression”. Lastly, Fisher's exact test was used to test two networks for enrichment for causal cognition genes. The cognition gene list was assembled by querying the Mouse Genome Database⁵⁴ under the general category, MP:0002063. Next, genes in the two subcategories, enhanced learning (MP:0012314) and enhanced conditioning behavior (MP:0012316), were removed to ensure that the final gene list reflected genes causal solely for decreased cognitive behaviors. Fisher's nominal p -values were reported since only two modules were tested. Odds ratios for all modules were plotted for the sake of comprehensiveness.

Probabilistic causal modeling to identify directed gene–gene relationships. In this analysis, we examined previously reported graphical models representing gene–gene relationships in the F2 mouse population⁴². Constructing probabilistic causal models of gene expression data have been extensively detailed elsewhere⁵⁵. Briefly, these probabilistic causal models are directed acyclic graphs in which each node represents gene expression of a gene, and the relationship between genes is captured by the conditional probability between nodes⁵⁶. Monte Carlo Markov Chain (MCMC) simulation was used to reconstruct one thousand gene networks that fit the gene expression data and the fit of each reconstruction was assessed with Bayesian Information Criterion (BIC)—a metric conservative to overfitting since it gives a lower prior probability to more complex models (i.e. with more parameters). A consensus network was consequently calculated by keeping edges that appear in more than 30% of the 1000 reconstructions to balance sensitivity and specificity⁵⁷. To infer causality in Markov equivalent structures, we used previously calculated cis-eQTLs as priors⁴² and prevented genes without cis-eQTLs from being parents of genes with cis-eQTLs, under the assumption that they are primarily regulated by genetic variation.

Calculating key drivers of gene networks from probabilistic causal models. To calculate key drivers, we used a variant of module enrichment that utilizes the directedness of the probabilistic causal graphs^{55,58}. For each module of interest, the downstream neighborhood of each gene was calculated, and nodes whose downstream network is greater than two standard deviations above the mean are considered network drivers. This method classifies nodes based upon their downstream network and uses the principle that nodes with more downstream nodes likely have a greater influence on the gene expression of the module as a whole. Consequently, key drivers are those nodes with significantly larger downstream networks than the average node in the module.

Results

Characterizing functional gene networks associated with the NREM delta power. In a previous study, two hundred eighty-three (283) behavioral, affective, and neurophysiologic phenotypes were measured in a population of (C57BL/6J \times 129S1/SvImJ) F2 mice, revealing gene network pleiotropy and novel relationships

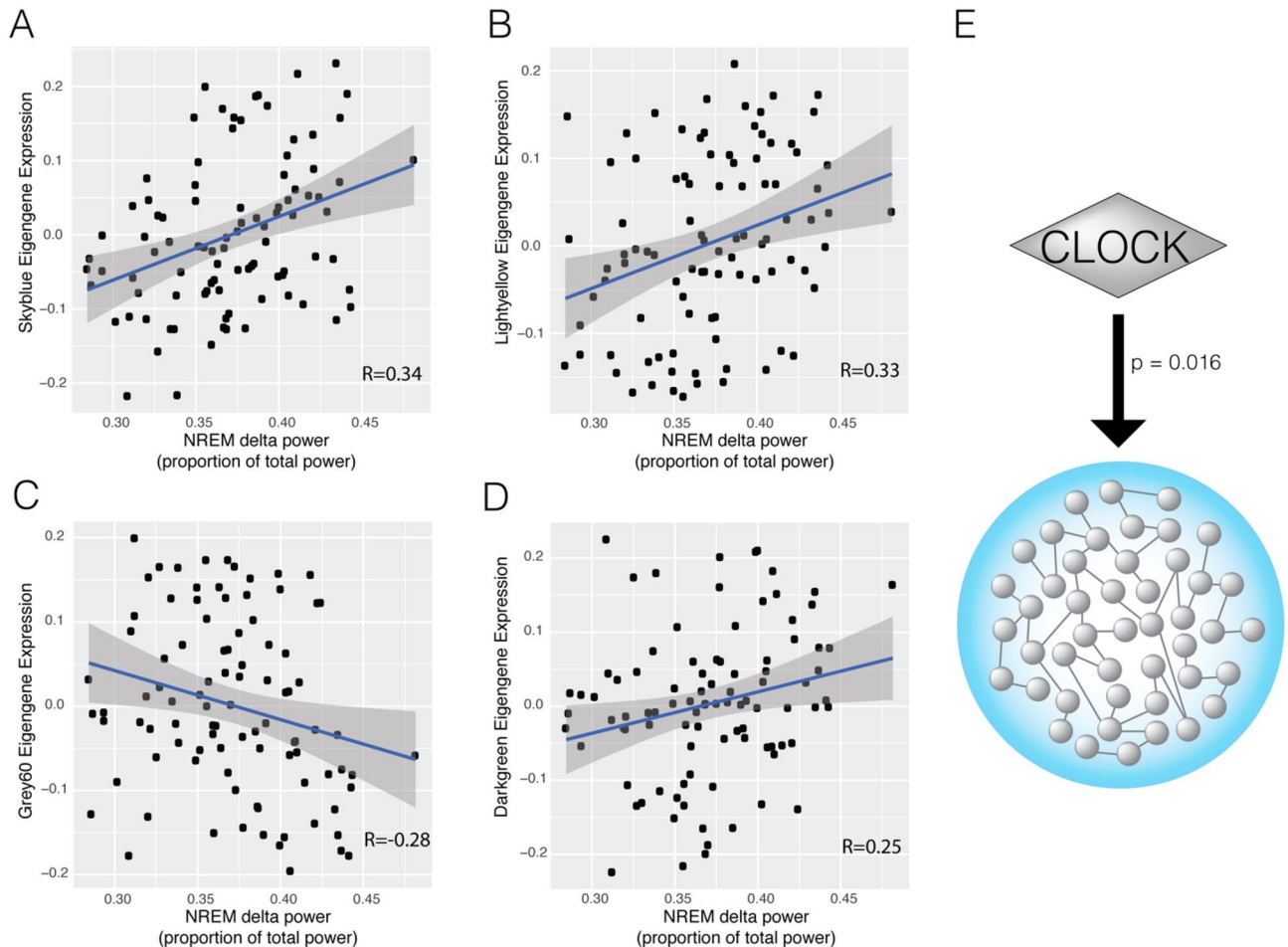


Figure 1. NREM delta power correlates with module eigengene expression of skyblue (A), lightyellow (B), grey60 (C), and darkgreen (D) modules. Correlation coefficients for each module-trait relationship is denoted in its respective panel. CLOCK knockdown directly modulates the expression of the skyblue module, with Fisher's p -value reported in the figure (E). Summary statistics for panels (A–D) were originally represented in Supplementary Table 3 in PMID: 30050989.

between sleep–wake and behavioral traits⁴². We used these data to specifically investigate cortical gene networks associated with NREM delta power. Notably, these gene networks were reconstructed in cortical regions rostral to the areas where the EEG leads were placed. Although the regulation of NREM delta power and NREM homeostasis has a strong local component^{59,60}, it also involves many brain areas. Particularly, prefrontal cortex in both humans and mice plays an important role in regulating NREM delta waves recorded in other cortical regions^{39,61}. Four cortical gene networks were correlated with NREM delta power during baseline conditions (Fig. 1A–D): darkgreen ($r_s=0.25$, $P=0.02$, FWER=0.03), grey60 ($r_s=-0.28$, $P=0.006$, FWER=0.01), lightyellow ($r_s=0.33$, $P=0.001$, FWER=0.005), and skyblue ($r_s=0.34$, $P=7.6 \times 10^{-4}$, FWER=0.003). Skyblue showed the strongest association with NREM delta power and was previously implicated in both sleep deprivation and major depressive disorder⁴². This network includes circadian clock genes and clock-regulator genes *Per1*, *Per2*, *Bhlhe40*, and *Klf10*. Further analyses of ChIP-seq and gene expression data demonstrated that skyblue genes are targeted by CLOCK ($P=9.2 \times 10^{-10}$, $Q=7.6 \times 10^{-8}$) and their expression is strongly modulated by CLOCK knockdown ($P=1.58 \times 10^{-3}$, $Q=0.016$)⁶² (Fig. 1E). These results indicate that NREM delta power is correlated with four cortical gene networks and most strongly associated with the expression of a CLOCK-dependent cortical gene network.

Identifying NREM delta power gene networks associated with AD progression and tauopathy.

Reduced NREM slow wave activity was associated with AV-1451 tau levels in PET scans of Brodmann area (BA) 10³⁵. Previous experiments have measured histopathologic markers and BA10 gene expression in patients with clinical symptoms across the AD spectrum, including patients with no symptoms at all⁴¹. These experiments indirectly captured information regarding the development of AD pathogenesis and the gene networks involved. We reasoned that the association between human NREM slow wave activity and AD tauopathy may be reflected by shared underlying gene networks. To investigate this relationship, we compared gene networks associated with NREM delta power to those networks previously associated with AD progression. This analysis showed that sixteen of the twenty-three (70%) cortical modules in the F2 mouse population are

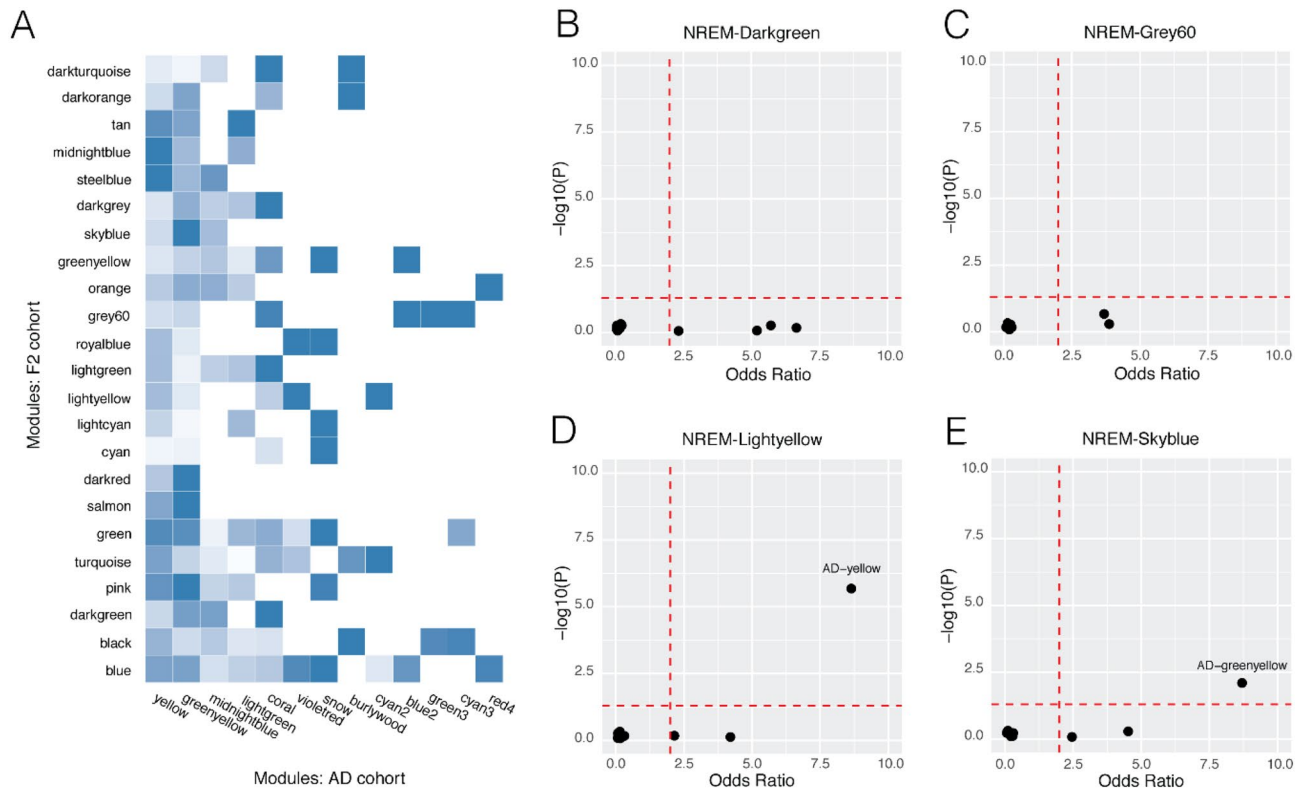


Figure 2. Sleep-relevant gene networks were compared directly with AD-relevant gene networks to determine network conservation (A). The darkness of the blue is proportional to the Fisher's exact test odds ratio, with darker colors suggesting greater overrepresentation. Each NREM network enrichment is represented on an independent plot (B–E). NREM network enrichment for each AD-relevant network is represented by an individual point on the plot. The odds ratio is plotted on the x-axis and the $-\log_{10}$ of the Fisher's P value plotted along the y-axis. (B–E) NREM networks were considered conserved in the AD cohort when odds ratio > 2 and $P < 0.05$, represented by the dotted red lines respectively.

conserved in the AD cohort. Of the modules associated with NREM delta power, lightyellow and skyblue are overrepresented in at least one AD gene network (Fig. 2A–E).

Next, we investigated if the AD modules associated with NREM delta power were also correlated with levels of tauopathy. NREM-lightyellow was overrepresented in AD-yellow ($P = 2.7 \times 10^{-6}$, OR = 8.5 [4.1–16.8]), a gene network previously correlated with Braak score ($r = -0.29$, $P = 0.022$) and neurofibrillary tangle density ($r = -0.33$, $P = 0.008$). NREM-skyblue was overrepresented in AD-greenyellow ($P = 0.01$, OR = 8.5 [2.9–21]), which was primarily associated neurofibrillary tangle density ($R = -0.25$, $P = 0.046$). Both AD-yellow and AD-greenyellow are strongly enriched for known Alzheimer's disease genes ($P = 1.8 \times 10^{-4}$ and $P = 1.1 \times 10^{-3}$, respectively), further confirming their relevance to AD pathogenesis. AD-yellow is composed of genes primarily related to synaptic function ($P = 1.6 \times 10^{-21}$), long-term potentiation ($P = 8.8 \times 10^{-5}$), and circadian entrainment ($P = 2.7 \times 10^{-3}$), while AD-greenyellow is primarily associated with mitochondrial biology ($P = 1.9 \times 10^{-14}$) and metal ion binding ($P = 5.3 \times 10^{-3}$). These results indicate that two cortical gene networks are associated with NREM delta power and AD tauopathy.

Gene networks associated with delta power are modulated in non-demented individuals with neuropathological evidence of Alzheimer's disease.

Gene networks relevant to NREM delta power and Alzheimer's-related tauopathy may indicate aspects of the molecular pathogenesis evident in the early stages of Alzheimer's disease. We have provided indirect evidence of this by identifying NREM delta power gene networks relevant to AD progression. To validate this finding, we examined an independent cohort of non-demented individuals with evidence of Alzheimer's neuropathology (NDAD)⁶³. These NDAD data were generated from BA10 and BA11 using laser capture microdissection and compared to an age- and sex-matched cohort that was both histopathologically and clinically normal. These data provide a useful approximation for examining the relevance of NREM delta power gene networks in preclinical AD. We reasoned that the two NREM delta power gene networks would be strongly modulated in NDAD if these networks reflected the early molecular stage of Alzheimer's disease. Our analysis showed that genes differentially expressed in NDAD compared to their histopathologically and clinically normal controls (Benjamini–Hochberg FDR < 0.1) are overrepresented in NREM-lightyellow and NREM-skyblue. NDAD strongly modulates both NREM-lightyellow ($P = 0.005$, Odds Ratio = 3.6 [1.4–9.3]) and NREM-skyblue ($P = 3.4 \times 10^{-7}$, Odds Ratio = 7.5 [3.2–19.6]), differentially regulating approximately 22% of the NREM-lightyellow gene network and 59% of NREM-skyblue (Fig. 3).

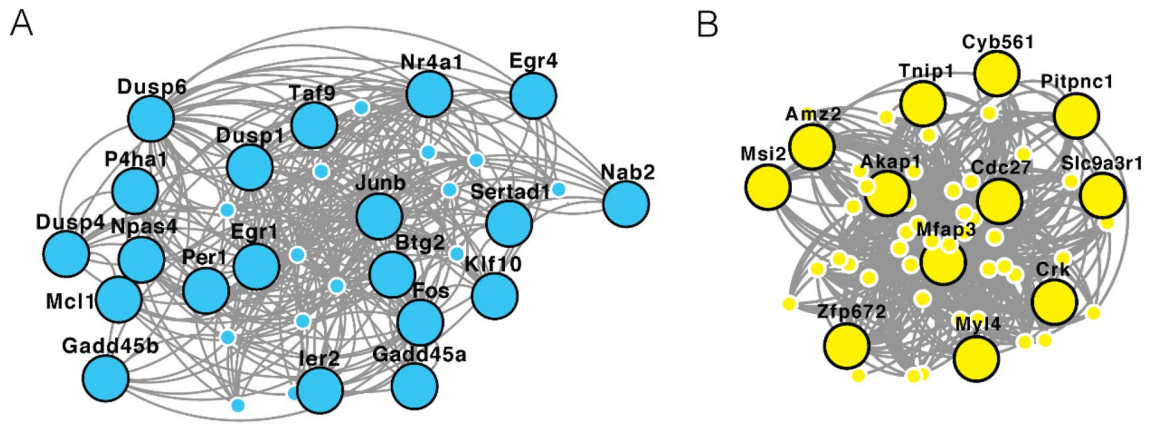


Figure 3. Coexpression networks for skyblue (A) and lightyellow (B) were calculated using weighted gene coexpression network analysis. The large, labeled nodes are the module members that are modulated in non-demented individuals with Alzheimer's neuropathology. Network images generated using Cytoscape v3.7.2 (<https://cytoscape.org/>).

Integrative probabilistic causal models reveal gene network regulators associated with delta power and preclinical AD.

Probabilistic causal models offer an effective strategy to integrate genetic variant and gene expression data and predict directional relationships between genes⁵⁷. These models identify highly reproducible gene–gene relationships⁶⁴, and empirical validation has demonstrated that these methods can successfully identify causal regulators in gene networks across a variety of species^{57,65–67}, including humans with late-onset Alzheimer's disease^{58,68,69}. Given the phenotypic and transcriptional relationship between NREM and AD, we reasoned that key drivers of NREM-skyblue and NREM-lightyellow may contribute to preclinical AD pathogenesis and AD progression. Previous work investigated the intramodular and intermodular gene–gene regulation in this F2 mouse population⁴². This earlier analysis independently highlighted the gene–gene relationships and key drivers in one of our networks-of-interest, NREM-skyblue, because it is strongly modulated by sleep deprivation and major depression. *Arc* and *Egr2* were identified as key regulators of NREM-skyblue (Fig. 4A) and intermodular analysis demonstrated that it was upstream of a number of other functional pathways, including genes involved in respiratory transport chain, TCA cycle, mitochondrial function, and synaptic processing⁴². In the present analysis, we also specifically report the gene–gene regulatory network of NREM-lightyellow. Analysis of the directed gene–gene network in NREM-lightyellow revealed that *Mrpl55*, a mitochondrial ribosomal protein, was its primary regulator, upstream of genes altered in NDAD (Fig. 4A). These key drivers highlight the importance in preclinical AD of core components of bioenergetics and synaptic processing.

Examining NREM delta power networks for known causal genes in learning, memory, and conditioning.

Lastly, we reasoned that these networks may causally influence cognitive behaviors. We projected genes known to cause abnormalities in learning, memory, and conditioning in mouse models⁵⁴ onto NREM-lightyellow and NREM-skyblue to determine if these networks may play a causal role in cognition. NREM-skyblue is strongly overrepresented with genes known to alter cognitive traits [P=0.002, OR=5.1 (1.7–12.6)] (Fig. 4B). These genes include the subnetwork key regulator, *Arc*, but known cognition-related genes are found throughout the downstream networks of both *Arc* and *Egr2*. On the other hand, we did not find such causal evidence for NREM-lightyellow [P>0.05, OR=0.89 (0.11–3.4)] (Fig. 4B), implying that its role in NDAD is not strongly causal or not captured by existing evidence for causal genes in cognition. These analyses provide corroborative evidence for the causal role of NREM-skyblue in NDAD and equivocal evidence for NREM-lightyellow.

Discussion

Alzheimer's disease progresses over the course of many years, culminating in pathologic changes to memory, executive function, behavior, and personality. Our analyses leverage insights from genetic, transcriptional, neurophysiologic, and histopathologic data from human and mouse cohorts to investigate the common transcriptional elements shared between AD progression and NREM delta power. These experiments were designed as a molecular corroboration of a recent study that showed reduced NREM slow-wave activity in predominantly cognitive normal patients with tauopathy. We showed that two gene networks, including a CLOCK-dependent circadian network, are positively correlated with NREM delta power in a population of (C57BL/6J) × 129S1/SvImJ) F2 mice and negatively correlated with tauopathy—estimated by neuropathologic measures of the Braak score and neurofibrillary tangle density—in a human AD population modeling AD progression. We confirmed that these two gene networks are particularly relevant for preclinical AD by demonstrating that they are strongly modulated in non-demented patients with neuropathologic features of AD. Lastly, probabilistic causal modeling identified *Arc*, *Egr2*, and *Mrpl55* as key drivers of the network associated with preclinical AD, and further analysis confirmed that both *Arc*- and *Egr2*-related subnetworks include genes that alter memory and learning.

Network drivers possibly play an important role in the preclinical stages of AD. *Arc* is a critical regulator of synaptic memory^{70–73} and is required for activity-dependent generation of Aβ⁷⁴. Genetic variants of *Arc* confer genetic susceptibility to AD in Han Chinese⁷⁵ and neuroprotection in other populations⁷⁶. Our evidence points

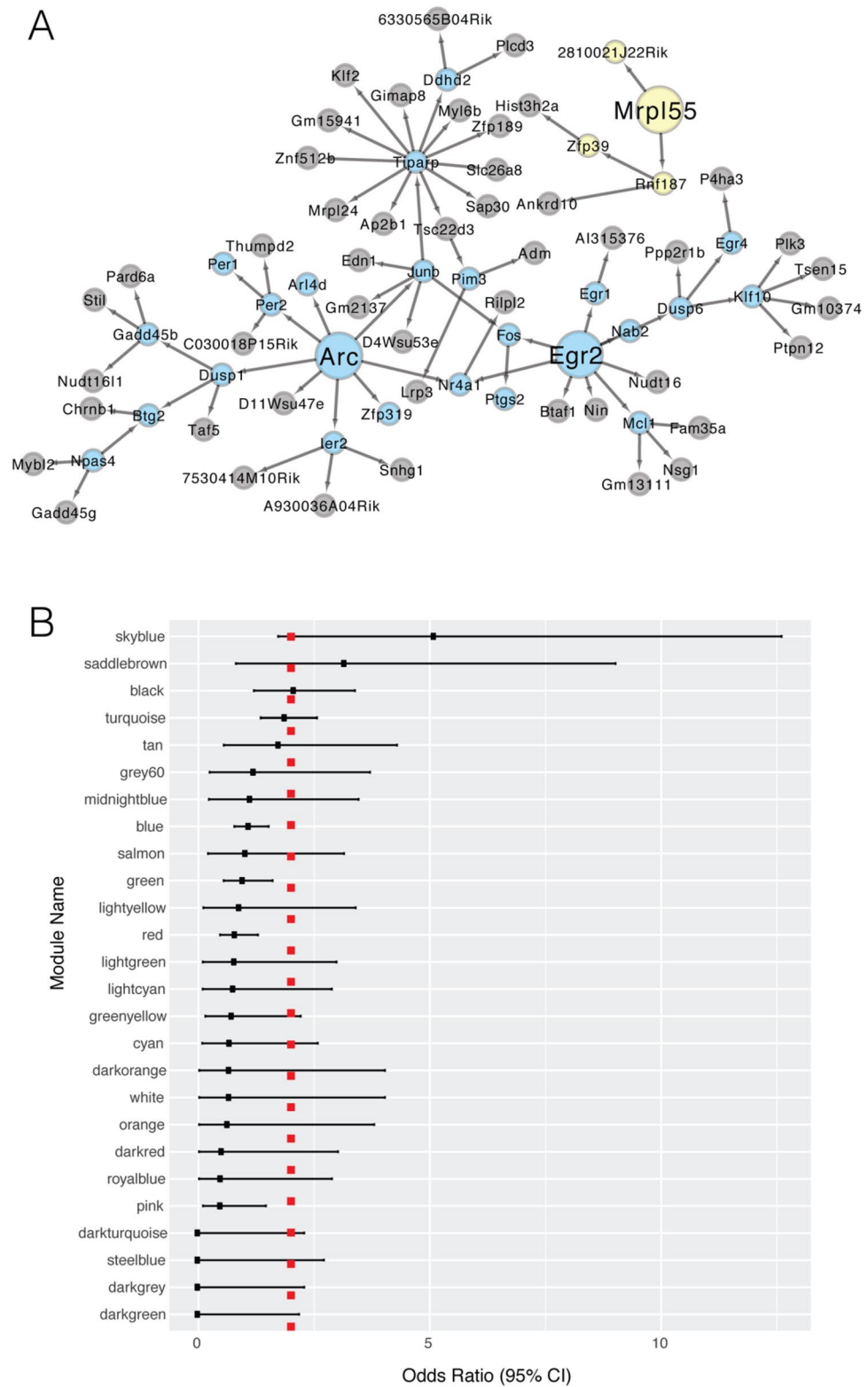


Figure 4. These direct acyclic graphs represent a subset of the transcriptome-wide gene–gene relationships involving skyblue and lightyellow genes (A). Skyblue and lightyellow nodes are highlighted by their respective colors, while all other genes, irrespective of their modules, are pictured in grey. The larger nodes represent the key regulators. Our panel (A) was generated using Cytoscape v3.7.2 (<https://cytoscape.org/>) and depicts a Bayesian network that was partially reproduced in Fig. 6B of PMID: 30050989. The Fisher’s odds ratio and 95% confidence interval are pictured (B), representing each network’s enrichment for genes known to cause abnormalities in learning, memory, and executive function. Networks are sorted from highest to lowest odds ratio. Dotted red line represents an odds ratio of 2.

to its role in AD progression and NREM slow wave activity disruption. These results are consistent with our previous study, demonstrating that *Arc* is a probable upstream regulator of genes known to causally influence affective behaviors and sleep patterns⁴², two phenotypes commonly dysregulated in preclinical AD. *Egr2* also has a well-documented role in learning and memory⁷⁷, and has also been associated with A β plaque-associated microglia activation⁷⁸. In our directed networks, *Egr2* is immediately upstream of *Egr1*, a transcriptional activator of BACE-1⁷⁹ and regulator of acetylcholinesterase⁸⁰. *Egr2* is also binds to and transcriptionally regulates *VGF*^{81,82}, a potential CSF biomarker for AD progression⁸³ and a recently discovered causal network mediator of Alzheimer's disease whose overexpression decreases cortical tau⁶⁹. Further, *Egr2*, as well as *Arc*, are strongly downregulated by medium chain triglycerides⁸⁴, which some evidence indicates may generally improve cognition and possibly treat APOE-negative Alzheimer's disease^{85,86}. Though there is strong evidence for the mechanisms of *Arc* and *Egr2*, there is very little experimental evidence for the biological function or pathologic role of *Mrp155*, pointing to a potential new target for future investigation.

Our approach indirectly characterizes the molecular features of the AD prodromal phase. Understanding molecular changes in preclinical AD may reveal novel aspects of pathogenesis masked by data collected in late stage cohorts and offer new therapeutic approaches. The present study had the advantage of directly validating networks associated with NREM delta power and AD progression in a cohort of non-demented individuals with AD-related pathology, a reasonable surrogate for preclinical AD. By integrating data across multiple cohorts, pre-clinical molecular networks can be reasonably inferred, partially mitigating the difficulty of measuring molecular changes in preclinical AD cohorts across time.

Alzheimer's disease and sleep deprivation have a bidirectional relationship, increasing the risk of one another and likely leading to a pathologic feedback loop¹⁹. Psychiatric symptoms, including depression, are also associated with neuropathologic features of Alzheimer's disease^{9,87}, possibly increasing the risk of neurodegeneration^{6,7}. Our systematic approach prioritized gene networks linked to tauopathy and NREM delta power, however, this analysis recovered a network that has been previously associated with sleep deprivation and major depressive disorder⁴². Interestingly, the network exhibits features of the bidirectional relationship between sleep and AD. Sleep deprivation strongly modulates this network's expression, while its key drivers regulate genes known to causally influence sleep-wake cycles. These findings suggest that many varied features of AD emergence and progression converge onto these cortical networks, which may serve as focal points for the complex, dynamic, and intertwined symptomatology.

Our method has several limitations that should be considered as readers interpret our results. Firstly, the transcriptome measurements in the mouse and AD cohorts were derived from bulk tissue. These data are opaque to high-resolution gene expression information, including single-cell and cell-type specific variation. A notable exception in our analyses were those data generated in non-demented individuals using laser capture microdissection. These experiments targeted specific neurons and complement our analysis of bulk tissue with higher resolution data. Another limitation is that our analysis integrates data from mouse and human brain regions that are functionally homologous to a certain degree, but have important anatomical and histological differences. Cortical homology between rodents and primates is debated, and important distinctions are often made between histological and functional homology^{88,89}. The convergence of results in our experiments, despite differences between species, argues obliquely for the utility of our approach, but the interpretation of our findings should include an understanding of the inherent structural and functional differences between mouse and human prefrontal cortex. Lastly, our findings reveal molecular networks consistent with features of preclinical AD, including gene targets predicted by our modeling, and these findings were indirectly corroborated by analyzing numerous independent data sets and investigating previous experimental literature. Nevertheless, careful prospective experimental study is required to dissect the functional role of key drivers in AD prodrome and pathogenesis.

Received: 19 March 2020; Accepted: 27 July 2020

Published online: 08 April 2021

References

- McDade, E. *et al.* Longitudinal cognitive and biomarker changes in dominantly inherited Alzheimer disease. *Neurology* **91**, e1295–e1306 (2018).
- Bateman, R. J. *et al.* Clinical and biomarker changes in dominantly inherited Alzheimer's disease. *N. Engl. J. Med.* **367**, 795–804 (2012).
- Morris, J. C. & Price, J. L. Pathologic correlates of nondemented aging, mild cognitive impairment, and early-stage Alzheimer's disease. *J. Mol. Neurosci.* **17**, 101 (2001).
- Lancôt, K. L. *et al.* Neuropsychiatric signs and symptoms of Alzheimer's disease: New treatment paradigms. *Alzheimer's Dementia Transl. Res. Clin. Interv.* **3**, 440–449 (2017).
- Zhao, Q.-F. *et al.* The prevalence of neuropsychiatric symptoms in Alzheimer's disease: Systematic review and meta-analysis. *J. Affect Disord.* **190**, 264–271 (2016).
- Masters, M. C., Morris, J. C. & Roe, C. M. "Noncognitive" symptoms of early Alzheimer disease. *Neurology* **84**, 617–622 (2015).
- Sacuiu, S. *et al.* Chronic depressive symptomatology in mild cognitive impairment is associated with frontal atrophy rate which hastens conversion to Alzheimer dementia. *Am. J. Geriatr. Psychiatry* **24**, 126–135 (2016).
- Donovan, N. J. *et al.* Longitudinal association of amyloid beta and anxious-depressive symptoms in cognitively normal older adults. *Am. J. Psychiatry* **175**, 530–537 (2018).
- Donovan, N. J. *et al.* Association of higher cortical amyloid burden with loneliness in cognitively normal older adults. *JAMA Psychiatry* <https://doi.org/10.1001/jamapsychiatry.2016.2657> (2016).
- Rosenberg, P. B. Loneliness as a marker of brain amyloid burden and preclinical Alzheimer disease. *JAMA Psychiatry* <https://doi.org/10.1001/jamapsychiatry.2016.2688> (2016).
- Holth, J. K. *et al.* The sleep-wake cycle regulates brain interstitial fluid tau in mice and CSF tau in humans. *Science* **363**, eaav546 (2019).

12. Carvalho, D. Z. *et al.* Association of excessive daytime sleepiness with longitudinal β -amyloid accumulation in elderly persons without dementia. *JAMA Neurol.* <https://doi.org/10.1001/jamaneurol.2018.0049> (2018).
13. Shi, L. *et al.* Sleep disturbances increase the risk of dementia: A systematic review and meta-analysis. *Sleep Med. Rev.* **40**, 4–16 (2018).
14. Shokri-Kojori, E. *et al.* β -Amyloid accumulation in the human brain after one night of sleep deprivation. *Proc. Natl. Acad. Sci.* **115**, 201721694 (2018).
15. Sprecher, K. E. *et al.* Poor sleep is associated with CSF biomarkers of amyloid pathology in cognitively normal adults. *Neurology* **89**, 445–453 (2017).
16. Benedict, C. *et al.* Self-reported sleep disturbance is associated with Alzheimer's disease risk in men. *Alzheimer's Dementia* **11**, 1090–1097 (2015).
17. Peter-Derex, L., Yammine, P., Bastuji, H. & Croisile, B. Sleep and Alzheimer's disease. *Sleep Med. Rev.* **19**, 29–38 (2015).
18. Spira, A. P. *et al.* Self-reported sleep and β -amyloid deposition in community-dwelling older adults. *JAMA Neurol.* **70**, 1537–1543 (2013).
19. Ju, Y.-E.S., Lucey, B. P. & Holtzman, D. M. Sleep and Alzheimer disease pathology—A bidirectional relationship. *Nat. Rev. Neurol.* **10**, 115–119 (2013).
20. Sterniczuk, R., Theou, O., Rusak, B. & Rockwood, K. Sleep disturbance is associated with incident dementia and mortality. *Curr. Alzheimer Res.* **10**, 767–775 (2013).
21. Kang, J.-E. *et al.* Amyloid- β dynamics are regulated by orexin and the sleep-wake cycle. *Science* **326**, 1005–1007 (2009).
22. Anderson, K. E. *et al.* Clinical management of neuropsychiatric symptoms of Huntington disease: Expert-based consensus guidelines on agitation, anxiety, apathy, psychosis and sleep disorders. *J. Huntington's Dis.* **7**, 355–366 (2018).
23. Epping, E. A. *et al.* Longitudinal psychiatric symptoms in prodromal Huntington's disease: A decade of data. *Am. J. Psychiatry* **173**, 184–192 (2016).
24. Lazar, A. S. *et al.* Sleep deficits but no metabolic deficits in premanifest Huntington's disease. *Ann. Neurol.* **78**, 630–648 (2015).
25. Videnovic, A., Lazar, A. S., Barker, R. A. & Overeem, S. 'The clocks that time us'—Circadian rhythms in neurodegenerative disorders. *Nat. Rev. Neurol.* **10**, 683–693 (2014).
26. Gustafsson, H., Nordström, A. & Nordström, P. Depression and subsequent risk of Parkinson disease. *Neurology* **84**, 2422–2429 (2015).
27. Schrag, A., Horsfall, L., Walters, K., Noyce, A. & Petersen, I. Prediagnostic presentations of Parkinson's disease in primary care: A case-control study. *Lancet Neurol.* **14**, 57–64 (2015).
28. Howell, M. J. & Schenck, C. H. Rapid eye movement sleep behavior disorder and neurodegenerative disease. *JAMA Neurol.* **72**, 707–712 (2015).
29. Gehrman, P., Gooneratne, N. S., Brewster, G. S., Richards, K. C. & Karlawish, J. Impact of Alzheimer disease patients' sleep disturbances on their caregivers. *Geriatr. Nurs.* **39**, 60–65 (2018).
30. Liu, S. *et al.* Caregiver burden and prevalence of depression, anxiety and sleep disturbances in Alzheimer's disease caregivers in China. *J. Clin. Nurs.* **26**, 1291–1300 (2017).
31. Caceres, B. A. *et al.* Family caregivers of patients with frontotemporal dementia: An integrative review. *Int. J. Nurs. Stud.* **55**, 71–84 (2016).
32. D'Aoust, R. F., Brewster, G. & Rowe, M. A. Depression in informal caregivers of persons with dementia. *Int. J. Older People Nurs.* **10**, 14–26 (2015).
33. Musiek, E. S. & Holtzman, D. M. Mechanisms linking circadian clocks, sleep, and neurodegeneration. *Science* **354**, 1004–1008 (2016).
34. Miyata, S. *et al.* Poor sleep quality impairs cognitive performance in older adults. *J. Sleep Res.* **22**, 535–541 (2013).
35. Lucey, B. P. *et al.* Reduced non-rapid eye movement sleep is associated with tau pathology in early Alzheimer's disease. *Sci. Transl. Med.* **11**, eaau6550 (2019).
36. Lucey, B. P. *et al.* Effect of sleep on overnight cerebrospinal fluid amyloid β kinetics. *Ann. Neurol.* **83**, 197–204 (2018).
37. Musiek, E. S. *et al.* Circadian rest-activity pattern changes in aging and preclinical Alzheimer disease. *JAMA Neurol.* <https://doi.org/10.1001/jamaneurol.2017.4719> (2018).
38. Mander, B. A. *et al.* β -amyloid disrupts human NREM slow waves and related hippocampus-dependent memory consolidation. *Nat. Neurosci.* **18**, 1051–1057 (2015).
39. Mander, B. A. *et al.* Prefrontal atrophy, disrupted NREM slow waves and impaired hippocampal-dependent memory in aging. *Nat. Neurosci.* **16**, 357 (2013).
40. Ju, Y.-E.S. *et al.* Slow wave sleep disruption increases cerebrospinal fluid amyloid- β levels. *Brain* **140**, 2104–2111 (2017).
41. Wang, M. *et al.* Integrative network analysis of nineteen brain regions identifies molecular signatures and networks underlying selective regional vulnerability to Alzheimer's disease. *Genome Med.* **8**, 104 (2016).
42. Scarpa, J. R. *et al.* Cross-species systems analysis identifies gene networks differentially altered by sleep loss and depression. *Sci. Adv.* **4**, eaat1294 (2018).
43. Diessler, S. *et al.* A systems genetics resource and analysis of sleep regulation in the mouse. *PLoS Biol.* **16**, e2005750 (2018).
44. Vassalli, A. & Franken, P. Hypocretin (orexin) is critical in sustaining theta/gamma-rich waking behaviors that drive sleep need. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E5464–E5473 (2017).
45. Franken, P., Chollet, D. & Tafti, M. The homeostatic regulation of sleep need is under genetic control. *J. Neurosci.* **21**, 2610–2621 (2001).
46. Haroutunian, V., Katsel, P. & Schmeidler, J. Transcriptional vulnerability of brain regions in Alzheimer's disease and dementia. *Neurobiol. Aging* **30**, 561–573 (2009).
47. Zhang, B. & Horvath, S. A general framework for weighted gene co-expression network analysis. *Stat. Appl. Genet. Mol.* **4**, Article 17 (2005).
48. Ravasz, E., Somera, A. L., Mongru, D. A., Oltvai, Z. N. & Barabási, A. L. Hierarchical organization of modularity in metabolic networks. *Science* **297**, 1551–1555 (2002).
49. Shannon, P. *et al.* Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
50. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37**, 1–13 (2009).
51. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44 (2009).
52. Chen, E. Y. *et al.* Enrichr: Interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinform.* **14**, 128 (2013).
53. Alhopuro, P. *et al.* Mutations in the circadian gene CLOCK in colorectal cancer. *Mol. Cancer Res.* **8**, 952–960 (2010).
54. Blake, J. A. *et al.* The Mouse Genome Database: Integration of and access to knowledge about the laboratory mouse. *Nucleic Acids Res.* **42**, D810–D817 (2014).
55. Jiang, P. *et al.* A systems approach identifies networks and genes linking sleep and stress: Implications for neuropsychiatric disorders. *Cell Rep.* **11**, 835–848 (2015).
56. Zhu, J. *et al.* An integrative genomics approach to the reconstruction of gene networks in segregating populations. *Cytogenet. Genome Res.* **105**, 363–374 (2004).

57. Zhu, J. *et al.* Increasing the power to detect causal associations by combining genotypic and expression data in segregating populations. *PLoS Comput. Biol.* **3**, e69 (2007).
58. Zhang, B. *et al.* Integrated systems approach identifies genetic nodes and networks in late-onset Alzheimer's disease. *Cell* **153**, 707–720 (2013).
59. Rattenborg, N. C., Lima, S. L. & Lesku, J. A. Sleep locally, act globally. *Neuroscience* **18**, 533–546 (2012).
60. Krueger, J. M. & Tononi, G. Local use-dependent sleep; synthesis of the new paradigm. *Curr. Top. Med. Chem.* **11**, 2490–2492 (2011).
61. Dort, C. J. V., Baghdoyan, H. A. & Lydic, R. Adenosine A(1) and A(2A) receptors in mouse prefrontal cortex modulate acetylcholine release and behavioral arousal. *J. Neurosci. Off. J. Soc. Neurosci.* **29**, 871–881 (2009).
62. Lachmann, A. *et al.* ChEA: Transcription factor regulation inferred from integrating genome-wide ChIP-X experiments. *Bioinformatics* **26**, 2438–2444 (2010).
63. Liang, W. S. *et al.* Neuronal gene expression in non-demented individuals with intermediate Alzheimer's Disease neuropathology. *Neurobiol. Aging* **31**, 549–566 (2010).
64. Cohain, A. *et al.* Exploring the reproducibility of probabilistic causal molecular network models. *Pac. Sympos. Biocomput.* **22**, 120–131 (2017).
65. Chen, Y. *et al.* Variations in DNA elucidate molecular networks that cause disease. *Nature* **452**, 429–435 (2008).
66. Yang, X. *et al.* Validation of candidate causal genes for obesity that affect shared metabolic pathways and networks. *Nat. Genet.* **41**, 415–423 (2009).
67. Schadt, E. E. *et al.* An integrative genomics approach to infer causal associations between gene expression and disease. *Nat. Genet.* **37**, 710–717 (2005).
68. Readhead, B. *et al.* Multiscale analysis of independent Alzheimer's cohorts finds disruption of molecular, genetic, and clinical networks by human herpesvirus. *Neuron* **99**, 64–82.e7 (2018).
69. Beckmann, N. D. *et al.* Multiscale causal network models of Alzheimer's disease identify VGF as a key regulator of disease. *bioRxiv* <https://doi.org/10.1101/458430> (2018).
70. Korb, E. & Finkbeiner, S. Arc in synaptic plasticity: From gene to behavior. *Trends Neurosci.* **34**, 591–598 (2011).
71. Béique, J.-C., Na, Y., Kuhl, D., Worley, P. F. & Huganir, R. L. Arc-dependent synapse-specific homeostatic plasticity. *Proc. Natl. Acad. Sci.* **108**, 816–821 (2011).
72. Shepherd, J. D. & Bear, M. F. New views of Arc, a master regulator of synaptic plasticity. *Nat. Neurosci.* **14**, 279 (2011).
73. Peebles, C. L. *et al.* Arc regulates spine morphology and maintains network stability in vivo. *Proc. Natl. Acad. Sci.* **107**, 18173–18178 (2010).
74. Wu, J. *et al.* Arc/Arg3.1 regulates an endosomal pathway essential for activity-dependent β -amyloid generation. *Cell* **147**, 615–628 (2011).
75. Initiative, A. D. N. *et al.* The arc gene confers genetic susceptibility to Alzheimer's disease in Han Chinese. *Mol. Neurobiol.* **55**, 1217–1226 (2018).
76. Kettunen, P. *et al.* A novel ARC polymorphism is associated with reduced risk of Alzheimer's disease. *Alzheimer's Dementia* **6**, S196 (2010).
77. Poirier, R. *et al.* Paradoxical role of an Egr transcription factor family member, Egr2/Krox20, in learning and memory. *Front. Behav. Neurosci.* **1**, 6 (2007).
78. Yin, Z. *et al.* Immune hyperreactivity of A β plaque-associated microglia in Alzheimer's disease. *Neurobiol. Aging* **55**, 115–122 (2017).
79. Qin, X., Wang, Y. & Paudel, H. K. Early growth response 1 (Egr-1) is a transcriptional activator of β -secretase 1 (BACE-1) in the brain. *J. Biol. Chem.* **291**, 22276–22287 (2016).
80. Hu, Y. *et al.* Early growth response-1 regulates acetylcholinesterase and its relation with the course of Alzheimer's disease. *Brain Pathol.* <https://doi.org/10.1111/bpa.12688> (2019).
81. Matys, V. *et al.* TRANSFAC and its module TRANSCompel: Transcriptional gene regulation in eukaryotes. *Nucleic Acids Res.* **34**, D108–D110 (2006).
82. Matys, V. *et al.* TRANSFAC: Transcriptional regulation, from patterns to profiles. *Nucleic Acids Res.* **31**, 374–378 (2003).
83. Hendrickson, R. C. *et al.* High resolution discovery proteomics reveals candidate disease progression markers of Alzheimer's disease in human cerebrospinal fluid. *PLoS ONE* **10**, e0135365 (2015).
84. Wang, D. & Mitchell, E. S. Cognition and synaptic-plasticity related changes in aged rats supplemented with 8- and 10-carbon medium chain triglycerides. *PLoS ONE* **11**, e0160159 (2016).
85. Sharma, A., Bemis, M. & Desilets, A. R. Role of medium chain triglycerides (Axona) in the treatment of mild to moderate Alzheimer's disease. *Am. J. Alzheimer's Dis. Other Dementias* **29**, 409–414 (2014).
86. Henderson, S. T. & Poirier, J. Pharmacogenetic analysis of the effects of polymorphisms in APOE, IDE and IL1B on a ketone body based therapeutic on cognition in mild to moderate Alzheimer's disease; a randomized, double-blind, placebo-controlled study. *BMC Med. Genet.* **12**, 137 (2011).
87. Ehrenberg, A. J. *et al.* Neuropathologic correlates of psychiatric symptoms in Alzheimer's disease. *J. Alzheimer's Dis.* **66**, 1–12 (2018).
88. Laubach, M., Amarante, L. M., Swanson, K. & White, S. R. What, if anything, is rodent prefrontal cortex?. *Eneuro.* <https://doi.org/10.1523/ENEURO.0315-18.2018> (2018).
89. Carlén, M. What constitutes the prefrontal cortex?. *Science* **358**, 478–482 (2017).

Acknowledgements

The views, opinions, and/or findings contained in this work are those of the authors and should not be interpreted as representing the official views or policies, either expressed or implied, of the Defense Advanced Research Projects Agency or the Department of Defense.

Author contributions

J.R.S. and P.J. wrote the main manuscript text. J.R.S., P.J., V.D.G. performed all analyses. J.R.S., A.K., M.H.V., F.W.T. conceived of the experimental design. All authors edited and reviewed the final manuscript and figures.

Funding

This work was supported by a research grant from Merck Research Laboratories, by the Defense Advanced Research Projects Agency, and by the U.S. Army Research Office (contract no. W911NF-10-1006). J.R.S. was supported by the National Institute of Mental Health of the NIH (F30MH106293), and V.D.G. was supported by the National Heart, Lung, and Blood Institute of the NIH (T32HL007909).

Competing interests

F.W.T. has received grants from Merck & Co. Inc. A.K was a former employee of Merck & Co. Inc. All other authors declare that they have no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-86255-6>.

Correspondence and requests for materials should be addressed to A.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021