


RESEARCH

Open Access



# Artificial intelligence framework identifies candidate targets for drug repurposing in Alzheimer's disease

Jiansong Fang<sup>1†</sup>, Pengyue Zhang<sup>2†</sup>, Quan Wang<sup>3,4†</sup>, Chien-Wei Chiang<sup>5</sup>, Yadi Zhou<sup>1</sup>, Yuan Hou<sup>1</sup>, Jielin Xu<sup>1</sup>, Rui Chen<sup>3,4</sup>, Bin Zhang<sup>1</sup>, Stephen J. Lewis<sup>6</sup>, James B. Leverenz<sup>7,8</sup>, Andrew A. Pieper<sup>9,10,11,12,13</sup>, Bingshan Li<sup>3,4\*</sup>, Lang Li<sup>5\*</sup>, Jeffrey Cummings<sup>14</sup> and Feixiong Cheng<sup>1,7,15\*</sup> 

## Abstract

**Background:** Genome-wide association studies (GWAS) have identified numerous susceptibility loci for Alzheimer's disease (AD). However, utilizing GWAS and multi-omics data to identify high-confidence AD risk genes (ARGs) and druggable targets that can guide development of new therapeutics for patients suffering from AD has heretofore not been successful.

**Methods:** To address this critical problem in the field, we have developed a network-based artificial intelligence framework that is capable of integrating multi-omics data along with human protein–protein interactome networks to accurately infer accurate drug targets impacted by GWAS-identified variants to identify new therapeutics. When applied to AD, this approach integrates GWAS findings, multi-omics data from brain samples of AD patients and AD transgenic animal models, drug–target networks, and the human protein–protein interactome, along with large-scale patient database validation and in vitro mechanistic observations in human microglia cells.

**Results:** Through this approach, we identified 103 ARGs validated by various levels of pathobiological evidence in AD. Via network-based prediction and population-based validation, we then showed that three drugs (pioglitazone, febuxostat, and atenolol) are significantly associated with decreased risk of AD compared with matched control populations. Pioglitazone usage is significantly associated with decreased risk of AD (hazard ratio (HR) = 0.916, 95% confidence interval [CI] 0.861–0.974,  $P = 0.005$ ) in a retrospective case-control validation. Pioglitazone is a peroxisome proliferator-activated receptor (PPAR) agonist used to treat type 2 diabetes, and propensity score matching cohort studies confirmed its association with reduced risk of AD in comparison to glipizide (HR = 0.921, 95% CI 0.862–0.984,  $P = 0.0159$ ), an insulin secretagogue that is also used to treat type 2 diabetes. In vitro experiments showed that

\*Correspondence: bingshan.li@vanderbilt.edu; Lang.Li@osumc.edu; chengf@ccf.org

†Jiansong Fang, Pengyue Zhang and Quan Wang contributed equally to this work.

<sup>1</sup> Genomic Medicine Institute, Lerner Research Institute, Cleveland Clinic, Cleveland, OH 44195, USA

<sup>3</sup> Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, TN 37212, USA

<sup>5</sup> Department of Biomedical Informatics, College of Medicine, Ohio State University, Columbus, OH 43210, USA

Full list of author information is available at the end of the article



pioglitazone downregulated glycogen synthase kinase 3 beta (GSK3 $\beta$ ) and cyclin-dependent kinase (CDK5) in human microglia cells, supporting a possible mechanism-of-action for its beneficial effect in AD.

**Conclusions:** In summary, we present an integrated, network-based artificial intelligence methodology to rapidly translate GWAS findings and multi-omics data to genotype-informed therapeutic discovery in AD.

**Keywords:** Alzheimer's disease, Drug repurposing, Genome-wide association studies (GWAS), Multi-omics, Network medicine, Pioglitazone

## Background

Alzheimer's disease (AD) is a chronic neurodegenerative disorder associated with progressive cognitive decline, extracellular amyloid plaques, intracellular neurofibrillary tangles, and neuronal death [1, 2]. AD and other dementias are an increasingly important global health burden, recently estimated to affect 43.8 million people worldwide [3]. Although genome-wide association studies (GWAS) have identified over 40 genome-wide significant susceptibility loci for AD [4–7], translating these findings into identification of high-confidence AD risk genes (ARGs) and potential therapies has eluded the field. Indeed, since Dr. Alois Alzheimer first described the condition in 1906, there are only five small-molecule drugs approved by the U.S. Food and Drug Administration (FDA) for treatment of AD: three cholinesterase inhibitors (donepezil, galantamine, and rivastigmine), one N-methyl-D-aspartate (NMDA) receptor antagonist (memantine), and one fixed combination of donepezil and memantine [8]. Aducanumab, a monoclonal antibody targeting aggregated beta-amyloid, is the first disease-modifying drug approved by U.S. FDA for Alzheimer's treatment in nearly 20 years; yet, its clinical efficacy is limited to a narrow segment of the AD continuum and potential side effects [1, 9].

The number of AD patients is expected to rise to 13.8 million by 2050 in the United States (U.S.) alone [10, 11], while the attrition rate for AD clinical trials (2002–2012) is estimated at 99.6% [12]. One possible explanation for why most candidate drugs fail in later-stage clinical trials is poor target selection. Broadly in disease, drug targets with genetic support have carried a high success rate among U.S. Food and Drug Administration (FDA)-approved therapies [13, 14]. However, this has not been the case with AD, and the translational application of multi-omics data such as GWAS for target identification and therapeutic development in AD remains challenging.

We recently demonstrated the utility of network-based methodologies for accelerating target identification and therapeutic discovery by exploiting multi-omics profiles from individual patients in multiple complex diseases, including cardiovascular disease [15], cancer [16], schizophrenia [17], and AD [18, 19]. We now posit that systematic identification of likely causal genes by incorporating GWAS findings and multi-omics profiles with human

interactome network models will also reveal disease-specific targets for genotype-informed therapeutic discovery in AD. This approach entails unique integration of the genome, transcriptome, proteome, and the human protein–protein interactome. In this study, we presented a network-based artificial intelligence (AI) framework that is capable of integrating multi-omics data along with human protein–protein interactome networks to accurately infer drug targets impacted by GWAS-identified variants to identify new therapeutics. Specifically, under the AI framework, we first apply a Bayesian algorithm to infer AD risk genes (termed ARGs) from AD GWAS loci via integrating multi-omics data and gene networks. Then repurposable drugs will be prioritized by quantifying the network proximity score [15, 16] of ARGs and drug targets in the human protein–protein interactome. Finally, we test the drug user's relationship with AD using large-scale, longitudinal patient data and further investigate drug's mechanism-of-action using in vitro mechanistic observations in human microglia cells (Fig. 1).

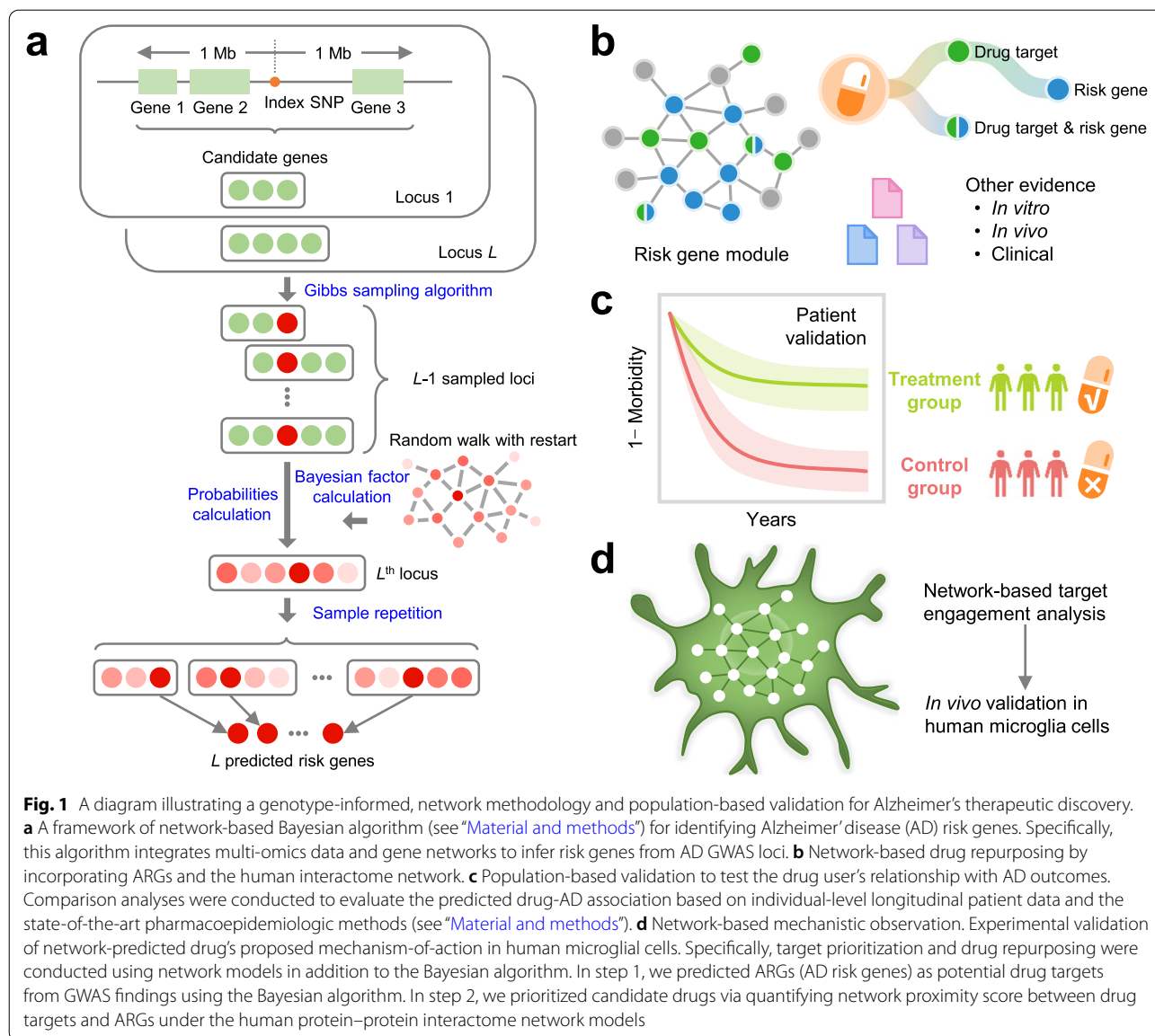
## Material and methods

### Collection of GWAS SNPs from large-scale studies

In this study, we assembled multiple single-nucleotide polymorphisms (SNPs) associated with AD from 15 large-scale GWAS studies in diverse population groups, conducted between 2007 and 2019 (Table S1). Some of the collected SNPs may represent the same genetic signal due to the use of overlapping samples across studies. To avoid this bias, we filtered the collected SNPs to remove redundant genetic signals (Supplementary Methods). To maximize genetic signals based on the omnigenic hypothesis [20], we adopted a loose threshold ( $P < 1 \times 10^{-5}$ ) to collect and filter AD SNPs and finally obtained 106 unique GWAS SNPs for downstream analyses.

### Construction of human protein–protein interactome

To build a comprehensive human protein–protein interactome, we assembled data from 15 common resources with multiple levels of experimental evidence (Supplementary Methods). Specifically, we focused on high-quality protein–protein interactions (PPIs) with the following five types of experimental data: (1) binary PPIs tested by high-throughput yeast-two-hybrid (Y2H)



systems; (2) kinase-substrate interactions by literature-derived low-throughput and high-throughput experiments; (3) literature-curated PPIs identified by affinity purification followed by mass spectrometry (AP-MS), Y2H, and by literature-derived low-throughput experiments, and protein three-dimensional structures; (4) signaling network by literature-derived low-throughput experiments; (5) protein complex data (see [Supplementary Methods](#)). The genes were mapped to their Entrez ID based on the National Center for Biotechnology Information (NCBI) database ([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)), and duplicated pairs were removed. Collectively, the integrated human interactome included 351,444 PPIs

connecting 17,706 unique proteins. More details are provided in our recent studies [15, 16].

### Collection of functional genomics data

We collected the distal regulatory element (DRE)-promoter links inferred from two studies. The first study was the capture Hi-C study of a lymphoblastoid cell line (GM12878) and we obtained 1,618,000 DRE-promoter links predicted from GM12878 [21]. The second dataset we used was from the Functional Annotation of the Mammalian Genome 5 (FANTOM5) project [22], in which cap analysis of gene expression (CAGE) technology was employed to infer enhancer-promoter links

across multiple human tissues. We downloaded FANTOM5 data and obtained 66,899 enhancer-promoter links [22].

#### Disease-associated genes

Open Targets database refers to a comprehensive platform for therapeutic target identification and validation [23]. We collected 527 AD-associated genes (Table S2) from the Open Targets database (accessed in September, 2019).

#### Seed genes with experimental evidence for Alzheimer's disease

We further collected 144 AD seed genes having either genetic, experimental, or functional evidence reported in large-scale GWAS studies, AD transgenic animal models, or human-derived samples (Table S2). These genes are involved in pathobiology of amyloidosis, tauopathy, or both, and genes characterizing other AD pathological hypotheses including neuroinflammation, vascular dementia, and other pathobiological pathways (Supplementary Methods).

#### Brain-specific gene expression

We downloaded RNA-Seq data (transcripts per million, TPM) of 31 tissues from GTEx V8 release (accessed on March 31, 2020, <https://www.gtexportal.org/home/>). We defined those genes with counts per million (CPM)  $\geq 0.5$  in over 90% of samples as tissue-expressed genes and the other genes as tissue-unexpressed. To quantify the expression significance of tissue-expressed gene  $i$  in tissue  $t$ , we calculated the average expression  $\langle E(i) \rangle$  and the standard deviation  $\delta_E(i)$  of a gene's expression across all tissues evaluated. The significance of gene expression in tissue  $t$  is defined as

$$z_E(i, t) = (E(i, t) - \langle E(i) \rangle) / \delta_E(i) \quad (1)$$

The details have been described in previous studies [15, 16].

#### Gene expression from single-cell/nucleus transcriptomics

We collected mouse single-cell/nucleus RNA sequencing (sc/snRNA-Seq) data in 5XFAD brain samples versus controls from two recent studies (GSE98969 and GSE140511) [24, 25]. We also collected human snRNA-Seq datasets on AD patient brain tissues from two publications [26, 27]. The first set of human snRNA-Seq data contains 10 frozen post-mortem human brain tissues from both entorhinal cortex (EC) and superior frontal gyrus (SFG) regions. This dataset has been deposited in the AD knowledge portal (<https://adknowledgeportal.synapse.org>, Synapse ID: syn21788402).

The raw data were deposited on Gene Expression Omnibus (GEO ID: GSE147528), which contains astrocytes, excitatory neurons, inhibitory neurons, and microglia cells [27]. We also assembled human snRNA-Seq data in AD cases versus controls with entorhinal cortex samples across six major brain cell types (GEO ID: GSE138852): microglia, astrocytes, neurons, oligodendrocytes, oligodendrocyte progenitor cells (OPCs), and endothelial cells [26].

The original sc/snRNA-Seq datasets were downloaded from the GEO database ([www.ncbi.nlm.nih.gov/geo/](http://www.ncbi.nlm.nih.gov/geo/)), and detailed information of these datasets is provided in Table S3. The analyses were completed with Seurat (v3.1.5), scran (v1.16.0), scater (v1.16.1) packages in R with steps complied with the original literature [24–27]. Data were normalized using a scaling factor of 10,000, and all differential gene expression analyses were conducted by function *FindMarkers* in the Seurat R package with parameter `test.use = "MAST"`. All mouse genes were further mapped to unique human-orthologous genes using the Mouse Genome Informatics (MGI) database (Eppig *et al.*, 2017). Details of processing of sc/snRNA-seq data and quality control are provided in Supplementary Methods and our recent study [18].

#### Gene expression from microarray

We collected human microarray data in AD cases versus healthy controls with human brain samples from two independent datasets (GSE29378 and GSE84422) [28, 29]. We also collected mouse microarray data from AD transgenic mouse vs. controls, including brain microglia of 5XFAD mice from 2 independent datasets (GSE65067 and GSE74615) [30, 31], and brain hippocampus of Tg4510 mice (GSE53480 and GSE57583) [32].

The original microarray datasets were obtained from Gene Expression Omnibus (<https://www.ncbi.nlm.nih.gov/geo/>). Detailed information of these 6 GEO datasets is provided in Table S3. All raw expression data were log2 transformed, and all samples were quantile normalized together. Probe IDs in each dataset were mapped to NCBI Entrez IDs, and probes mapping to multiple genome regions or without corresponding entrez IDs were deleted. The items were imported to R statistical processing environment using a LIMMA/Bioconductor package. All the mouse genes were further transferred into unique human-orthologous genes using the MGI database [33]. Genes with threshold fold change (FC)  $> 1.2$  were defined as exhibiting differential expression and prioritized as predicted AD risk genes.

#### Bulk RNA sequencing data

We collected 2 RNA-seq datasets from brain or brain microglia of 5XFAD mice [34]. In addition, we obtained



4 RNA-seq datasets from brain microglia of Tg4510 mice across different months [M] age (2M, 4M, 6M, and 8M) [35]. Differential expression analysis was performed using DESeq [36], while threshold for significance of differential expression was set to FDR < 0.05 using Benjamini-Hochberg's method. After mapping mouse genes to human-orthologous gene [33], we obtained 6 differentially expressed gene sets.

**Proteomic data in AD models**

In total, 10 proteomic datasets were assembled from 3 types of AD transgenic mouse models in two recent publications [37, 38]. The first study performed global quantitative proteomic analysis in hAPP and hAPP-PS1 mouse models at young (3 month [M]) and old ages (12 M) [37]. We obtained four sets of DEPs (hAPP\_3M, hAPP\_12M, hAPP-PS1\_3M, and hAPP-PS1\_12M) after merging the DEPs from different brain regions. The second study performed quantitative proteomics to uncover molecular and functional signatures in the hippocampus of three types of transgenic mice [38]. Two of these mouse lines, including ADLPAPT (4M, 7M, 10M) that carry three human transgenes (APP, PSEN1, and tau) and hAPP-PS1 (4M, 7M, 10M) mouse, were used in this study. After mapping mouse genes to human-orthologous gene [33], we obtained 10 sets of DEPs.

**Enrichment analysis**

Differentially expressed gene/protein (DEG/DEP) sets from multiple data sources were collected for enrichment analysis using Fisher's exact test. This included a total of 6 bulk RNA-seq datasets and 10 proteomic datasets from 4 types of AD transgenic mouse models, including 5XFAD, Tg4510, ADLPAPT, and hAPP (see [Supplementary Methods](#)).

**AD risk gene prediction**

We utilized a Bayesian model selection method, adapted from our recent work [17], to predict ARGs. Specifically, we collected at most 20 genes in the 2-Mb region centered at a GWAS index SNP as the candidates for that particular locus. Assigning  $L$  as the number of GWAS loci, and we then denoted a vector of genes with length  $L$ , each being from one of the  $L$  GWAS loci, as  $(X_1, \dots, X_L)$ , and termed it as candidate risk gene set (CRGS). Assigning  $N$  to represent the biological network, we then calculated  $P(X_1, \dots, X_L|N)$  with the goal to select a CRGS with maximum posterior probability. Computationally, it is not feasible to enumerate all possible gene combinations, and we therefore adopted a Gibbs sampling algorithm to transition the problem into a single-dimensional sampling procedure. For example, when sampling the risk gene from candidates at the  $L$ th locus, we assumed

that the risk genes at all other  $L-1$  loci had been selected, and the sampling probability for a gene at the  $L$ th locus was computed as conditional on the  $L-1$  risk genes, based on its closeness to other  $L-1$  risk genes in the network. For each candidate gene  $X_L$  at the  $L$ th locus, we assigned  $M_1$  to represent the event that  $X_L$  is the risk gene at locus  $L$ ,  $M_0$  represent the event that  $X_L$  is not the risk gene at locus  $L$ , and  $X_{-L}$  to represent all the selected risk genes in the other  $L-1$  loci. The Bayesian model selection can be depicted as

$$\frac{P(M_1|X_{-L}, N)}{P(M_0|X_{-L}, N)} = \frac{P(M_1) P(X_{-L}|M_1, N)}{P(M_0) P(X_{-L}|M_0, N)} \tag{2}$$

where  $\frac{P(X_{-L}|M_1, N)}{P(X_{-L}|M_0, N)}$  is a Bayesian Factor (BF) measuring the closeness between  $X_{-L}$  and  $X_L$  in network  $N$  and  $\frac{P(M_1)}{P(M_0)}$  is prior odds. The prior odds reflect the prior knowledge whether  $X_L$  is a risk gene or not and we assumed  $P(M_1) = P(M_0)$  in this study.

In regard to  $\frac{P(X_{-L}|M_1, N)}{P(X_{-L}|M_0, N)}$ , we adopted the random walk with restart (RWR) algorithm to calculate the BF. Starting from any node  $n_i$  in a predefined network  $N$ , the walker faces two options at each step: either moving to a direct neighbor with a probability  $1 - r$  or jumping back to  $n_i$  with a probability  $r$ . The fixed parameter  $r$  is called the restart probability in RWR, and  $r$  was set as 0.3 in this study [17]. Let  $W$  be the adjacency matrix that decides which neighbor to be moved to, and  $q_t$  be the reaching probability of all nodes at step  $t$ . The RWR algorithm is formalized as

$$q_{t+1} = (1 - r)Wq_t + rs_{n_i} \tag{3}$$

$s_{n_i}$  is a vector with the  $i$ th element as 1 and 0 for others, which means the starting node is  $n_i$ . Following the equation,  $q_t$  can be updated step by step until  $|q_{t+1} - q_t|^2 < T_{rwr}$ , where  $T_{rwr}$  is a predefined threshold. We set  $T_{rwr}$  as  $1 \times 10^{-6}$  [17]. The adjacency matrix  $W$  represents the distance between any two nodes in the network, and we adopted the same network and strategy in our previous work to calculate  $W$ . We calculated  $P(X_{-L}|M_1, N)$  based on  $W$ . We mapped  $X_L$  to the rows of  $W$  and  $X_{-L}$  to the columns of  $W$ , and obtained a vector with the same length as  $X_{-L}$ . The sum of the vector was calculated as  $P(X_{-L}|M_1, N)$ . In this study, we assumed  $P(X_{-L}|M_0, N)$  to be the same for all different candidate genes. Through the Bayesian model selection equation,

$$\frac{P(M_1|X_{-L}, N)}{P(M_0|X_{-L}, N)} = \frac{P(M_1) P(X_{-L}|M_1, N)}{P(M_0) P(X_{-L}|M_0, N)} \tag{4}$$

we obtained a value for each candidate genes at locus  $L$ . We used these values as sampling for Gibbs sampling to choose a risk gene for locus  $L$ . We then repeated the sampling across the remaining loci and iterated the

sampling process until convergence. Specially, in each round of Gibbs sampling, we calculated the sampling frequency for each candidate gene. The frequency was compared with that of the previous round, and if the sum of squares of frequency differences across all selected genes was smaller than a predefined threshold ( $1 \times 10^{-4}$  used in this study), then the sampling procedure was halted. Based on the sampling, we are able to assess the confidence of candidates being risk genes.

### Construction of drug-target network

We integrated six commonly used resources to collect high-quality physical drug-target interactions for FDA-approved drugs. We obtained biophysical drug-target interactions using reported binding affinity data: inhibition constant/potency ( $K_i$ ), dissociation constant ( $K_d$ ), median effective concentration ( $EC_{50}$ ), or median inhibitory concentration ( $IC_{50}$ )  $\leq 10 \mu\text{M}$ . First, we extracted the bioactivity data from the DrugBank database (v4.3) [39], the Therapeutic Target Database (TTD, v4.3.02) [40], and the PharmGKB database [41]. To improve data quality, we pooled only those items that satisfied the following four criteria: (i) binding affinities, including  $K_i$ ,  $K_d$ ,  $IC_{50}$ , or  $EC_{50}$ ,  $\leq 10 \mu\text{M}$ ; (ii) the target protein has a unique UniProt accession number; (iii) proteins marked as “reviewed” in the UniProt database; and (iv) proteins are from *Homo sapiens*. Totally, we constructed a drug-target network including 15,367 physical drug-target interactions (edges), which connected 1608 FDA-approved drug nodes and 2251 unique human target nodes (Table S4).

### Description of network proximity

Given the set of disease proteins ( $A$ ), the set of drug targets ( $B$ ), then the closest distance  $d_{AB}$  measured by the average shortest path length of all the nodes to the other module in the human protein-protein interactome can be defined as:

$$\langle d_{AB} \rangle = \frac{1}{\|A\| + \|B\|} \left( \sum_{a \in A} \min_{b \in B} d(a, b) + \sum_{b \in B} \min_{a \in A} d(a, b) \right) \quad (5)$$

where  $d(a, b)$  denotes to the shortest path length between protein  $a$  and drug target  $b$ .

To calculate the significance of the network distance between a given drug and disease module, we constructed a reference distance distribution corresponding to the expected distance between two randomly selected groups of proteins of the same size and degree distribution as the original disease proteins and drug targets in the network. This procedure was run 1000 times. The

mean  $\bar{d}$  and standard deviation ( $\sigma_d$ ) of the reference distribution were used to calculate a  $z$ -score ( $z_d$ ) by converting an observed (non-Euclidean) distance  $d$  to a normalized distance.

### Pharmacoepidemiologic validation

#### Patient cohort preparation

The pharmacoepidemiology study utilized the MarketScan Medicare Supplementary database from 2012 to 2017. The dataset included individual-level diagnosis codes, procedure codes, and pharmacy claims for 7.23 million U.S. older adults (i.e., age  $\geq 65$  to be eligible for Medicare benefits) per year, which represents approximately 14% of the 46 million retirees with Medicare benefits. Pharmacy prescriptions of pioglitazone, febuxostat, atenolol, nadolol, sotalol, and glipizide were identified by using RxNorm and National Drug Code (NDC). For a subject exposed to the aforementioned drugs, a drug episode is defined as the time between drug initiation and drug discontinuation. Specifically, drug initiation is defined as the first day of drug supply (i.e., first prescription date). Drug discontinuation is defined as the last day of drug supply (i.e., last prescription date + days of supply) accompanied by no drug supply for the next 60 days. Gaps of less than 60-day of drug supply were allowed within a drug episode. For example, the pioglitazone cohort included the first pioglitazone episode for each subject, as well as the glipizide cohort. Further, we excluded observations that started within 180 days of insurance enrollment. For the final cohorts, demographic variables including age, gender and geographical location were collected. Additionally, diagnoses of hypertension (HTN), type 2 diabetics (T2D), and coronary artery disease (CAD), defined by The International Statistical Classification of Diseases (ICD) 9/10 codes (Supplementary Methods, Table S5), before drug initiation, were collected. These variables were specifically selected to address potential confounding biases. Lastly, a control cohort was selected from patients not exposed to that

drug (i.e., pioglitazone). Specifically, non-exposures were matched to the exposures (ratio 1:4) by initiation time of the drug, enrollment history, age, and gender. The geographical location, diagnoses of HTN, T2D, and CAD were collected for the control cohort as well.

#### Outcome measurement

The outcome was time from drug initiation to diagnosis of AD, which was defined by using the ICD codes

(Supplementary Methods). For the control cohort, the corresponding drug (i.e., pioglitazone) episode's starting date was used as the starting time. For pioglitazone and glipizide cohorts, observations without diagnosis of AD were censored at the end of drug episodes. Observations without diagnosis of AD were censored at the corresponding pioglitazone episode's end date (Fig. S1).

### Propensity score estimation

We define Location = region of residence (i.e., North East, North Central, South, and West), T2D = type 2 diabetes, HT = hypertension, and CAD = coronary artery disease.

The propensity score of taking repurposing drug vs. a comparator drug was estimated by the following logistic regression model:

$$\text{logit}[\text{Propensity Score}] \sim \text{Intercept} + \text{Age} + \text{Gender} + \text{Location} + \text{T2D} + \text{HT} + \text{CAD}. \quad (6)$$

Stratified Cox models were used to compare the AD risks. For repurposing drug vs. comparator drug or control, the analyses were stratified ( $n$  strata = 10) by the estimated propensity score. The propensity score adjusted Cox model is

$$\text{log}[\text{Hazard}] \sim \text{Strata}[\text{log}(\text{Baseline Hazard})|\text{Propensity Score}] + 1[\text{Repurposing drug yes}]. \quad (7)$$

For repurposing drug vs. control, the analyses were stratified based on the subgroups defined by gender, T2D diagnose, HT diagnoses, and CAD diagnoses.

### Statistical analysis

Survival curves for time to AD were estimated using a Kaplan-Meier estimator. Additionally, propensity score stratified survival analysis was conducted to investigate the risk of AD between pioglitazone users and pioglitazone non-users, febraxostat users and febraxostat non-users, atenolol users and atenolol non-users, nadolol users and nadolol non-users, and sotalol users and sotalol non-users. In addition, we conducted new comparison studies to calculate the risk of AD between pioglitazone users and glipizide users. For each comparison, the propensity score of taking each drug was estimated by using a logistic regression model in which covariates included age, gender, geographical location, T2D diagnosis, HTN diagnosis, and CAD diagnoses. Furthermore, propensity score stratified Cox-proportional hazards models were used to conduct statistical inference for the hazard ratios (HR) of developing AD between cohorts.

### Experimental validation

#### Reagents

Pioglitazone was acquired from Topscience. Lipopolysaccharides (LPS) (Cat# L2880) and 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) were obtained from Sigma-Aldrich. Antibodies against Phospho-GSK3B-Y216 (Cat# AP0261), GSK3B (Cat# A2081), and CDK5 (Cat# A5730) were purchased from ABclonal Technology. CDK5-Phospho-Tyr15 (Cat# YP0380) was obtained from Immunoway (Plano, Texas, USA). All other reagents were purchased from Sigma-Aldrich unless otherwise specified.

#### Cell viability

Human microglia HMC3 cells were purchased from American Type Culture Collection (ATCC, Manassas, VA). Cell viability was detected by MTT method. In total,

5000 cells/well were plated in 96-well plates for 12 h, and then treated with pioglitazone for 48 h. After treatment, MTT solution was added to the cells to a final concentration of 1 mg/mL, and the mixture was allowed to incubate at 37 °C for 4 h. The supernatant was removed,

and precipitates were dissolved in DMSO. Absorbance was measured at 570 nm using a Synergy H1 microplate reader (BioTek Instruments, Winooski, VT, USA).

#### Western blot analysis

HMC3 cells were pre-treated with pioglitazone (3 μM or 10 μM) and DMSO (control vehicle), and followed with 1 μg/mL LPS for 30 min. Cells were harvested, washed with cold PBS, and then lysed with RIPA Lysis Buffer with 1% Protease Inhibitor (Cat# P8340, Sigma-Aldrich). Total protein concentrations were measured using a standard BCA protein assay kit (Bio-Rad, CA, USA), according to the manufacturer's manual. Samples were electrophoresed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), then blotted onto a polyvinylidene difluoride (PVDF; EMD Millipore, Darmstadt, Germany) membrane. After transferring, membranes were probed with specific primary antibodies (1:1000) at 4 °C overnight. Specific protein bands were detected using a chemiluminescence reagent after hybridization with a horseradish peroxidase (HRP)-conjugated secondary antibody (1:3000).

## Results

### Pipeline of the network-based artificial intelligence methodology

We utilized a Bayesian model selection method to predict ARGs [17], based on the assumption that likely causal risk genes are more densely connected with each other in a biological network (Fig. 1a). By applying this Bayesian model to the 106 AD GWAS loci after being filtered the redundant genetic signals from original 366 SNPs (“Material and methods” and Table S1), we predicted 103 ARGs after merging the overlapping genes across several different loci (Table S6). Meanwhile, we also predicted a set of local background genes (LBGs) as a negative control for the following analyses [17]. We validated our 103 ARGs using multi-omics data, including functional genomic characteristics and transcriptomics, as well as proteomic profiles generated from diverse AD transgenic mice models and AD patient brain samples.

### Multi-omics validation of network-predicted risk genes in AD

Recent studies showed that disease-associated proteins tend to cluster in the same neighborhood of the human protein–protein interactome, forming a disease module, a connected subgraph that contains molecular determinants of a disease [15, 16]. Disease modules are commonly used to represent the molecular determinants of disease pathobiology/physiology in a variety of human diseases, including AD [8]. We found that 103 ARGs formed significantly connected subgraphs (termed disease module) rather than being scattered randomly in the human protein–protein interactome (Supplementary Methods), consistent with previous disease module analyses that we demonstrated in other multiple complex diseases [15, 16]. Specifically, 68.0% of ARGs (70/103,  $P = 0.015$ , permutation test) form the largest connected subnetwork (disease module), in comparison to the same number of randomly selected genes with similar connectivity (degree) as the original seed genes in the human interactome (Fig. S2). This disease module (Fig. 2a) includes 128 PPIs (edges or links) connecting 70 unique genes (nodes). Network analysis revealed 14 genes with

connectivity higher than 5, the top five of which were *ESR1*, *PSMC5*, *MAPK1*, *PAK1*, and *NFKB1*. These same five genes have previously been implicated in AD [42–45]. For example, *ESR1* interacts with tau protein in vivo, and prevents glutamate excitotoxic injury by  $A\beta$  via estrogen signaling [42]. Gene expression analysis shows that *PSMC5* was significantly overexpressed in patients carrying apolipoprotein E- $\epsilon 4$  (*APOE4*) mutations in comparison to *APOE* wild-type group [43]. In summary, 103 predicted ARGs comprise a strong disease module in the human interactome.

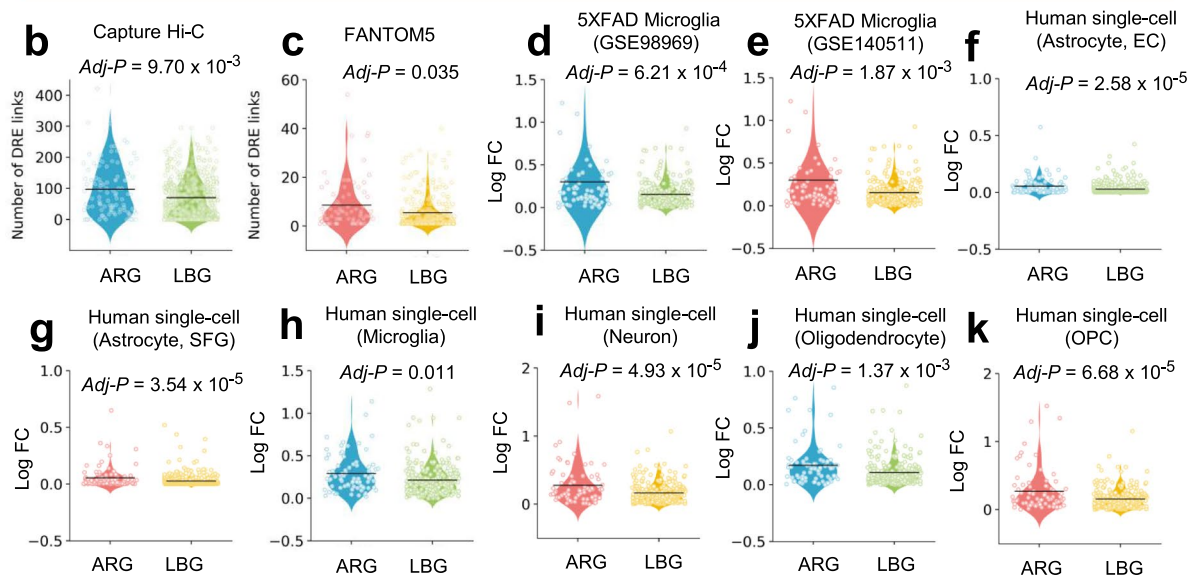
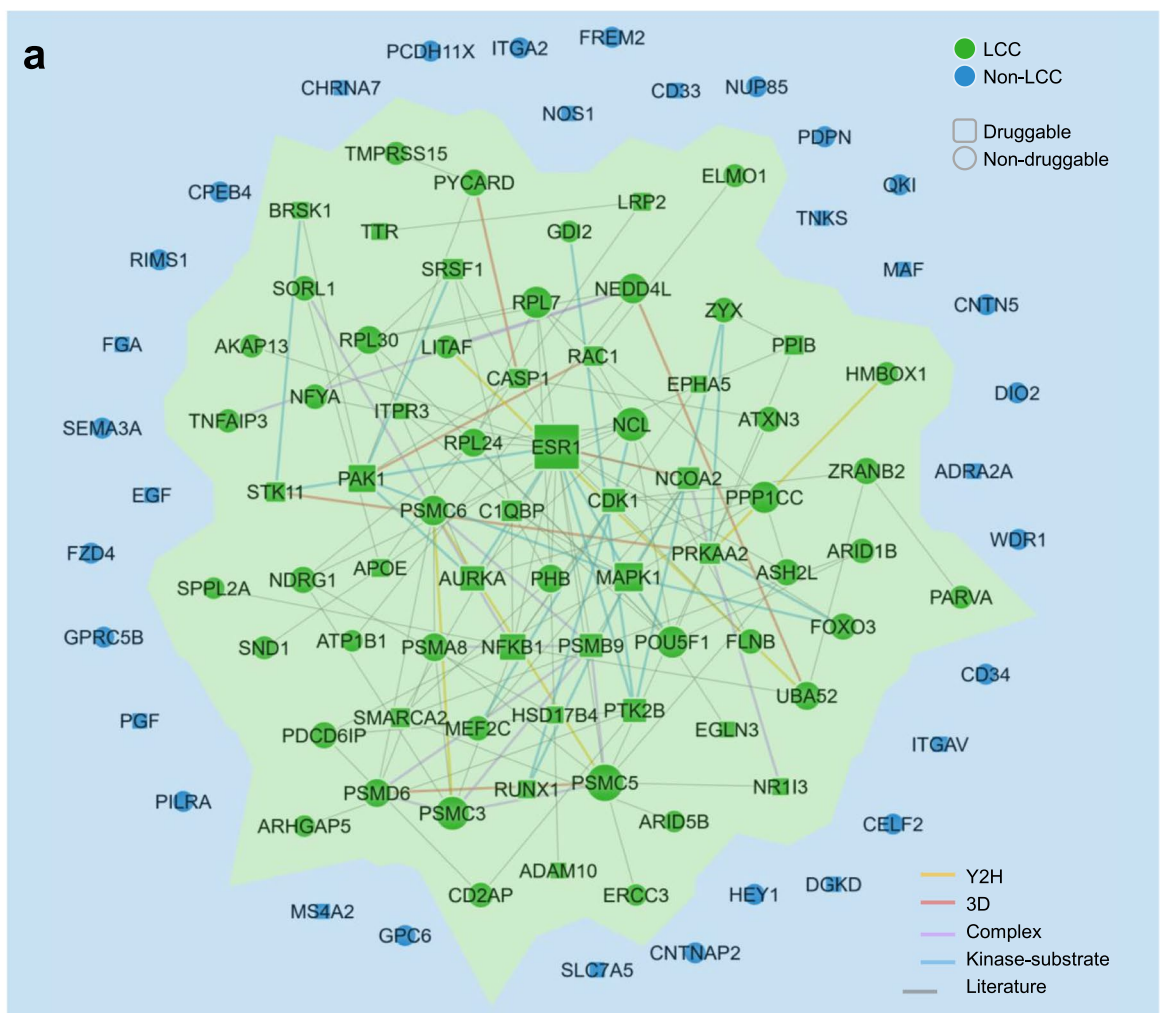
Because a majority of GWAS SNPs lie in noncoding region and exert their function by gene regulation [46], we explored the gene regulatory elements of ARGs by testing the hypothesis that the network-predicted risk genes capture more distal regulatory element (DRE)-promoter connections compared to 571 LBGs. We collected DRE-promoter connection data generated by two technologies: CAGE from FANTOM5 project and capture Hi-C (see “Material and methods”) [22, 47]. Through this, we found that the ARGs are indeed connected to more DREs in both capture Hi-C data ( $adj\text{-}P = 7.76 \times 10^{-3}$ , Fig. 2b) and FANTOM5 data ( $adj\text{-}P = 0.028$ , Fig. 2c).

We next investigated differential expression of ARGs under different pathobiology contexts of AD. We measured fold changes of gene expression levels of ARGs across 4 sc/snRNA-seq datasets (Table S3) compared to 571 LBGs. We found that ARGs were more likely to be differentially expressed in all sc/snRNA datasets (Fig. 2d–k). ARGs were more likely to be differentially expressed in (i) 5XFAD mouse brain microglial cells (Fig. 2d,e); (ii) human astrocyte cells of entorhinal cortex (Fig. 2f) and the superior frontal gyrus (Fig. 2g) from individuals spanning the neuropathological progression of AD; and (iii) a human single-cell atlas of entorhinal cortex from AD patients across four brain cell types: microglia (Fig. 2h), neuron (Fig. 2i), oligodendrocyte (Fig. 2j), and oligodendrocyte progenitor cell (OPC) (Fig. 2k). We further performed differentially expressed gene enrichment analysis for network-predicted ARGs in AD using bulk tissue expression data (Supplementary Methods). We collected bulk RNA-seq data from whole brain tissue

(See figure on next page.)

**Fig. 2** Network-based validation of predicted risk genes for Alzheimer’s disease (AD). **a** A subnetwork highlighting disease module formed by predicted AD risk genes (ARGs) in the human protein–protein interactome. This disease module includes 128 protein–protein interactions (PPIs) (edges or links) connecting 70 ARGs (nodes). Larger node size highlighting the high expression level in brain compared to other tissues. **b–k** Discovery of genomic features of 103 predicted ARGs implicated in AD. ARGs capture strong distal gene regulatory elements in Hi-C (**b**) and FANTOM5 data (**c**) compared to a set of local background genes (LBGs). **d–k** ARGs are more likely to be differentially expressed across 4 single-cell/nucleus RNA sequencing datasets (Table S3): **d, e** brain microglia cell of 5XFAD mouse model (GSE98969 [**d**] and GSE140511 [**e**]); **f, g** a human single-cell atlas (GSE147528) of entorhinal cortex (**f**) and the superior frontal gyrus (**g**) from individuals spanning the neuropathological progression of AD patient brain astrocyte cells; and a single-cell atlas (GSE138852) of entorhinal cortex from AD patients across four brain cell types: microglia [**h**], neuron [**i**], oligodendrocyte [**j**], oligodendrocyte progenitor cell (OPC) [**k**].  $P$  value was computed by one-tail  $T$ -test. Adjusted  $P$  value ( $adj\text{-}P$ ) was calculated based on the Benjamini–Hochberg approach. LCC: largest connected component; EC: entorhinal cortex; SFG: superior frontal gyrus





**Fig. 2** (See legend on previous page.)

or brain microglial cells from two common AD transgenic mouse models (5XFAD and Tg4510) and observed that ARGs were significantly differentially expressed in 5XFAD brain ( $P = 0.003$ ), 5XFAD microglial cells ( $P = 0.002$ ), and brain microglia of Tg4510 (Table S7). This suggests that our identified ARGs are potentially involved in diverse pathobiological pathways of AD.

We further inspected differentially expressed proteins encoded by 103 network-predicted ARGs across 10 published proteomics datasets (see [Supplementary Methods](#)) in AD. Herein, we evaluated 3 types of AD transgenic mouse models: (a) *hAPP* model containing *APP* transgene, (b) 5XFAD model harboring human transgenes for both *APP* and *PSEN1* mutations, and (c) ADLPAPT model carrying three human transgenes (*APP*, *PSEN1*, and *MAPT*). We found that products of ARGs were significantly differentially expressed in all 3 AD transgenic mouse models ( $P < 0.05$ , Fisher test, Table S7).

Collectively, we have thus shown that network-predicted ARGs are significantly involved in disease-related functional genomic, transcriptomic, and proteomic profiles, supporting their functional role as likely causal genes for AD.

#### Incorporation of AD multi-omics data to prioritize ARGs

We next turned to prioritize high-confidence ARGs by integrating multi-omics profiles. In total, we incorporated 8 criteria that can be categorized into 5 types of biological evidence: (1) brain-expression specificity ( $z$ -score) derived from GTEx database, (2) availability of supportive experimental evidence from the literatures and manually curated data from Open Targets database [23], (3) experimentally validated AD genes, (4) differential gene expression from microarrays, and (5) available drug targets. Figure 3 shows a global view of 103 ARGs that we validated by these multiple forms of biological evidence in AD. Among 103 ARGs, 89 genes (86.4% [89/103]) satisfy at least one criterion. To validate the remaining 14 ARGs without any omics evidence, we further collected significantly expressed proteins or genes from the most recent human AD brain proteomic or transcriptomic studies [26, 48–51]. We found that 7 were significantly expressed in five recent human AD brain proteomics or sc/snRNA-seq datasets (Table S8). In addition, among 103 ARGs, 13 ARGs have at least 5 types of AD-related evidence, including 8 well-known AD genes: *APOE*, *PTK2B*, *NOS1*, *MEF2C*, *SORL1*, *EPHA5*, *ADAM10*, and *CD33*. For the rest of the ARGs, all but *BRSK1* had corresponding published literature-derived evidence. For example, *PAK1* is a predicted risk gene with 6 criteria of biological evidence: high brain expression specificity ( $z$ -score = 1.01), supportive experimental evidence

from the literature, druggable target data, and differential expression in human brain of AD patients, microglial cells of 5XFAD mouse model, and brain hippocampus of a tau mouse model (Fig. 3 and Table S8). P21-activated kinase 1, encoded by the *PAK1* gene, has been implicated in AD [52], and recent studies have revealed that inactivation of *PAK1* obliterated social recognition without changing amyloid beta ( $A\beta$ )/tau pathology, and also exacerbated synaptic impairment and behavioral deficits in mouse models of AD [53, 54].

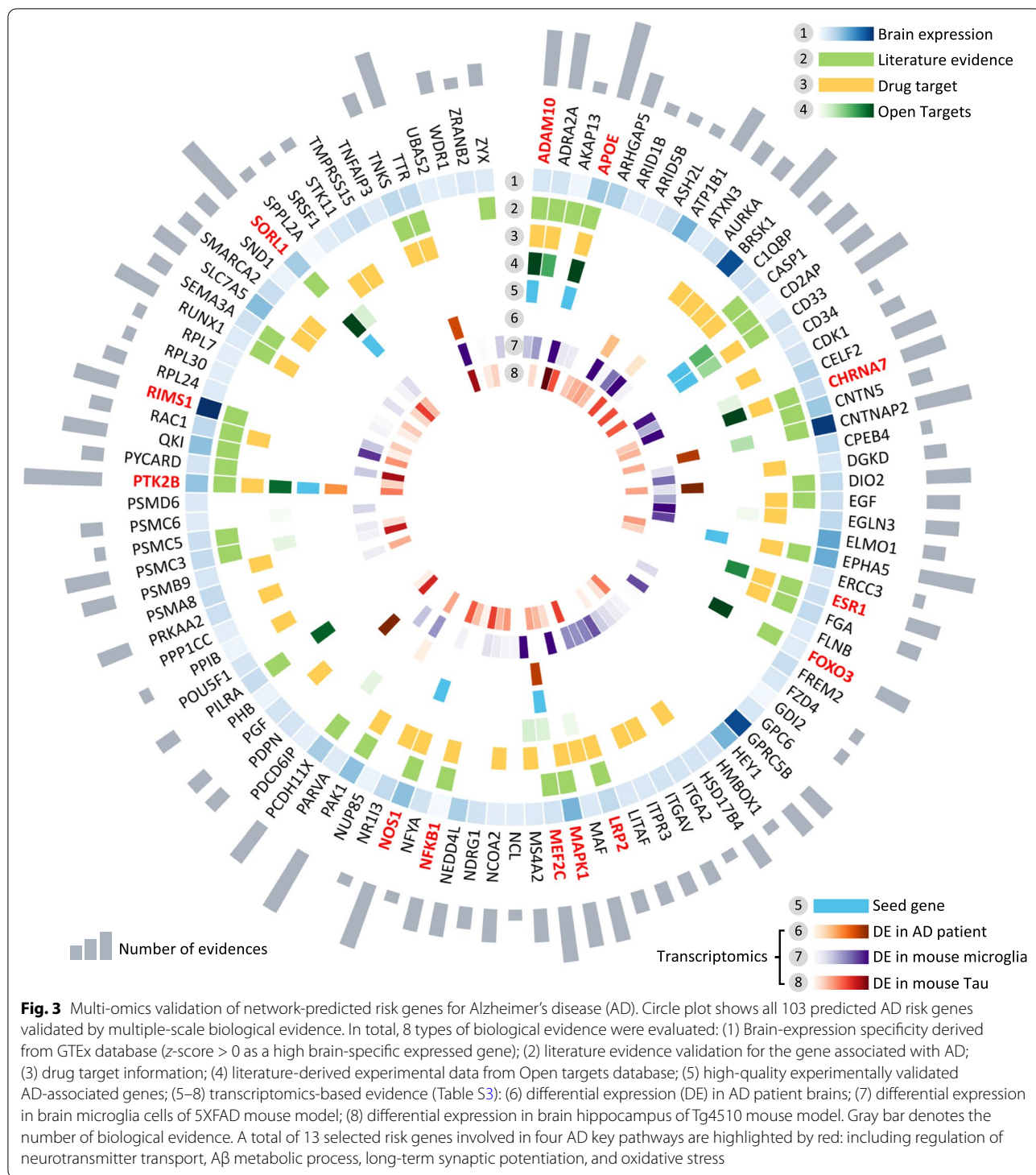
Among 103 ARGs, we selected 37 likely causal genes (Table S8) using subject matter expertise based on a combination of factors: (i) high brain-expression specificity, (ii) differential expression in multiple AD transgenic mouse models; (iii) strength of the network-based prediction, and (iv) availability of supportive experimental evidence. To advance disease understanding of network-predicted high-confidence risk genes, we performed biological pathway enrichment analysis using ClueGO plugin in Cytoscape (Table S9) [55]. We found 4 statistically significant biological pathways in AD and further discussed as below: (a) regulation of neurotransmitter transport, (b)  $A\beta$ -related biologic process, (c) long-term synaptic potentiation, and (d) oxidative stress (Table 1 and Table S10).

#### Neurotransmitter transport

Specifically, *MEF2C* and *RIMS1*, encoding myocyte-specific enhancer factor 2C and regulating synaptic membrane exocytosis protein 1, play key roles in neurotransmitter secretion and synaptic plasticity. *MEF2C* (rs254776) has been reported in several AD GWAS studies [56, 57], and we found significantly lower expression of *MEF2C* in AD brain (adj- $P = 0.010$ , one side Wilcoxon test, Fig. S3a). *RIMS1* is a newly predicted ARG, and a recent proteome study from human hippocampus revealed its overexpression in AD [58]. *RIMS1* is significantly overexpressed in 5XFAD mouse microglia (adj- $P = 0.036$ , one side Wilcoxon test) compared to controls (Fig. S3b).

#### Beta-amyloid-related biologic process

Five genes (*APOE*, *ADAM10*, *CHRNA7*, *SORL1*, and *LRP2*) are associated with beta-amyloid biologic process. Among them, *APOE*, *ADAM10*, and *SORL1* are well-known AD risk genes, validated by large-scale genetic studies and preclinical studies [6, 59, 60]. For example, *APP* cleavage by *ADAM10* will produce an APP-derived fragment that is neuroprotective, sAPP $\alpha$  [61]. *CHRNA7* (neuronal acetylcholine receptor subunit alpha-7) and *LRP2* (low-density lipoprotein receptor-related protein 2) are two newly identified risk genes. There is a significantly lower expression level of *CHRNA7* in the Tg4510



mouse (adj-*P* = 3.64 × 10<sup>-3</sup>) compared to controls (Fig. S3e). *CHRNA7* binds to Aβ with a high affinity [62]. Finally, a previous study showed that the rs3755166 polymorphism within *LRP2* is associated with susceptibility to AD in the Chinese population [63].

**Long-term synaptic potentiation**

Mitogen-activated protein kinase (*MAPK1*) and *PTK2B* are two identified risk genes related to long-term synaptic potentiation. Mitogen-activated protein kinase 1, encoded by *MAPK1* gene, is highly expressed in brain

**Table 1** Network-predicted risk genes involved in four pathobiological pathways of Alzheimer's disease (AD)

Gene	Protein	Description
<b>Neurotransmitter transport</b>		
<i>MEF2C</i> <sup>a</sup>	Myocyte-specific enhancer factor 2C	MEF2C mRNA expression levels were correlated with AD pathology
<i>RIMS1</i>	Regulating synaptic membrane exocytosis protein 1	An altered protein expression in RIMS1 during AD pathology
<b>Beta-amyloid-related biologic process</b>		
<i>APOE</i> <sup>a</sup>	Apolipoprotein E	Affect A $\beta$ production, aggregation, and clearance
<i>CHRNA7</i>	Neuronal acetylcholine receptor subunit alpha-7	Bind to A $\beta$ with very high affinity, providing therapeutic insight into AD
<i>SORL1</i> <sup>a</sup>	Sortilin-related receptor	Reduce A $\beta$ generation by trafficking APP away from amyloidogenic cleavage sites
<i>ADAM10</i> <sup>a</sup>	Disintegrin and metalloproteinase domain-containing protein 10	Constitutive $\alpha$ -secretase in the process of amyloid- $\beta$ protein precursor (A $\beta$ PP) cleavage
<i>LRP2</i>	Low-density lipoprotein receptor-related protein 2	rs3755166 polymorphism within LRP2 gene is associated with susceptibility to AD in the Chinese population
<b>Long-term synaptic potentiation</b>		
<i>MAPK1</i>	Mitogen-activated protein kinase 1	Beta-amyloid activates the MAPK cascade via hippocampal CHRNA7
<i>PTK2B</i> <sup>a</sup>	Protein-tyrosine kinase 2-beta	An in vivo modulator and early marker of Tau pathology
<b>Oxidative stress</b>		
<i>FOXO3</i>	Forkhead box protein O3	Activate BCL2L11 and FASLG to promote neuronal death and aberrant A $\beta$ processing
<i>NOS1</i>	Nitric oxide synthase	Loss of endothelial NOS promotes p25 production and Tau phosphorylation
<i>NFKB1</i>	Nuclear factor NF-kappa-B p105 subunit	Involve in neuroinflammation, synaptic plasticity, learning, and memory implicated in AD
<i>ESR1</i>	Estrogen receptor	Interact with tau protein in vivo, and prevent glutamate excitotoxic injury by A $\beta$ through estrogen signaling mechanisms

<sup>a</sup> Genes have experimental or functional evidence reported in AD transgenic animal models or human-derived samples (see Table S2 and Supplementary Method). The detailed literature data are provided in Table S2

tissue ( $z$ -score = 1.42). The *MAPK1* cascade can be activated by A $\beta$  via alpha7 nicotinic acetylcholine receptors [64]. *PTK2B*, a well-known AD gene with high expression in brain ( $z$ -score = 0.90), was identified as an early marker and in vivo modulator of tau pathology [65], by mediating A $\beta$ -induced synaptic dysfunction and loss [66]. There is significantly lower expression of *PTK2B* in AD patient brain transcriptome ( $adj$ - $P$  =  $2.05 \times 10^{-5}$ ) compared to controls (Fig. S3g).

#### Oxidative stress

Oxidative stress is a prominent hypothesis in the pathogenesis of AD [67]. Here we found four network-predicted ARGs (*FOXO3*, *NOS1*, *NFKB1*, and *ESR1*) that were associated with regulation of oxidative stress. *FOXO3* encoding Forkhead box protein O3 transcription factor is a direct substrate of *CDK5*. *FOXO3* activates several genes (e.g. *BCL2L11* and *FASLG*) to promote

neuronal death and aberrant A $\beta$  processing [68]. Significantly lower expression of *FOXO3* was found in 5XFAD mouse microglia ( $adj$ - $P$  =  $6.90 \times 10^{-3}$ ) compared to controls (Fig. S3h). *NFKB1*, encoding transcription factor nuclear factor kappa B (NF- $\kappa$ B), is implicated in oxidative stress, synaptic plasticity, and learning and memory [69].

Taken together, these findings suggest that our network-predicted ARGs are involved in diverse pathobiological pathways of AD. However, experimental validations are warranted for several newly predicted ARGs.

#### Network-predicted ARGs are more likely to be drug targets

To date, most disease genes generated from GWAS findings are undruggable [70]. For example, a recent study revealed that none of approved and investigational AD drugs target products (proteins) of GWAS-derived genes in AD [71]. We examined whether network-predicted ARGs were more druggable compared to randomly



selected proteins from human protein-coding gene background. Based on drug-target networks from 6 commonly used resources (see “[Material and methods](#)”), we obtained 2866 potential druggable proteins for FDA-approved or clinically investigational drugs. Surprisingly, we found that 41 out of 103 predicted ARGs (39.8 %) are known druggable proteins, which is four-fold higher than druggable proteins ( $P = 9.25 \times 10^{-11}$ , Fisher test) in the genome-wide human protein-coding genome background. High druggability of network-predicted ARGs offers more candidate targets for therapeutic discovery (such as drug repurposing) in AD. For example, *ADRA2A*, one of the predicted ARGs, encodes adrenoceptor alpha 2A receptor. *ADRA2A* is a potential target of clozapine [72], an atypical antipsychotic drug. Long-term clozapine treatment reduces A $\beta$  deposition and improves cognitive impairment in an AD transgenic mice model [73]. *NR1I3*, encoding the nuclear receptor constitutive androstane receptor (CAR), is a potential drug target activated by the lipid-lowering drug simvastatin [74]. Simvastatin was reported to significantly reduce levels of A $\beta$  in vitro and in vivo [75, 76]. In summary, network-predicted ARGs showed higher druggability compared to traditional GWAS-based analysis approaches. We next examined opportunities for drug repurposing by integrating findings from ARGs with the human protein-protein interactome network.

#### ARGs offer candidate targets for Alzheimer's drug repurposing

We have identified that network-predicted ARGs are related to the known pathobiology of AD and offer potential druggable targets, prompting us to examine opportunities for AD therapeutic discovery. We hypothesized that for a drug with multiple targets to be beneficial for treating a disease, its target proteins should be within or in the immediate vicinity of the corresponding disease module (Fig. 2a) in the human interactome network. To examine the potential application of ARGs on AD drug repurposing, we applied a network proximity approach [15] to quantify the interplay between AD modules from ARGs and drug targets in the human interactome network. We used the cutoff of  $z$ -score ( $z < -1.5$ ) to select network-predicted repurposable drugs in AD. After exclusion of nutraceutical drugs, metal drugs, and radioactive diagnostic agents, 130 drug candidates were obtained. We then systematically retrieved the published anti-AD clinical, in vitro/in vivo reported data for the 130 predicted drugs. In total, 25 had corresponding preclinical or clinical evidence for potential application to AD (Table S11). Figure 4 shows the molecular mechanisms of the 25 predicted drug candidates with published experimental or clinical evidence for AD. These drugs are classified into 6

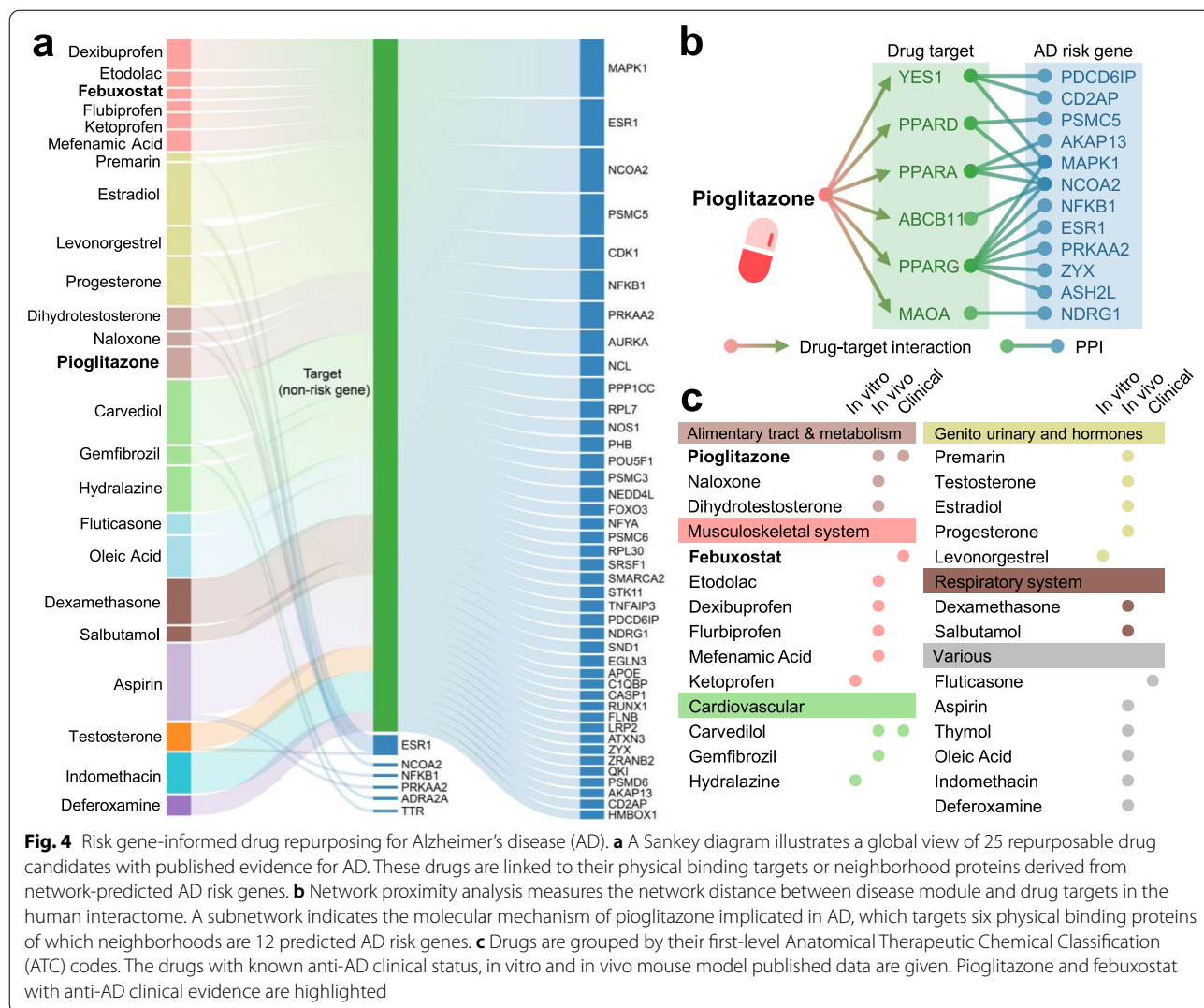
categories according to Anatomical Therapeutic Chemical classification (ATC) codes: musculoskeletal systems ( $n = 6$ ), genitourinary system and hormones ( $n = 5$ ), cardiovascular ( $n = 3$ ), alimentary tract and metabolism ( $n = 3$ ), respiratory system ( $n = 2$ ), and others ( $n = 6$ ).

Among them, we found 4 predicted drugs having known AD clinical evidence [77, 78], including febuxostat [79], pioglitazone (NCT02913664), carvedilol (NCT01354444), and fluticasone [77]. Febuxostat, a xanthine oxidase (XO) inhibitor approved for hyperuricemia, exerts a significant network proximity ( $z = -1.60$ ) with the ARGs. Pioglitazone, an FDA-approved drug for T2D, has a significant network proximity ( $z = -1.64$ ) with the ARGs. Figure 4 shows that pioglitazone targets six proteins by connecting with 12 neighborhoods of ARGs.

#### Network-predicted pioglitazone usage reduces risk of AD in patient data

To test both the febuxostat and pioglitazone users' relationships with AD outcomes using population-based validation, we conducted 2 rigorous retrospective case-control studies to compare AD risk by analyzing 7.23 million U.S. commercially insured individuals (“[Material and methods](#)”). These included the following: (i) pioglitazone ( $n = 101,650$ ,  $z = -1.64$  [network proximity score between ARGs and drug targets in the human interactome network]) vs. a matched control population (control,  $n = 402,488$ ), and (ii) febuxostat ( $n = 24,218$ ,  $z = -1.60$ ) vs. control ( $n = 95,192$ ). In order to identify more drug candidates with potential of reducing risk of AD, we conducted another 3 rigorous retrospective case-control studies for 3 antihypertensive agents with moderate  $z$ -score ( $z > -1.0$ ). These included the following: (iii) atenolol ( $n = 366,277$ ,  $z = -1.16$ ) vs. control ( $n = 1,449,815$ ); (iv) nadolol ( $n = 19,253$ ,  $z = -1.26$ ) vs. control ( $n = 76,136$ ); and (v) sotalol ( $n = 43,819$ ,  $z = -1.512$ ) vs. a control ( $n = 172,375$ ). Table 2 summarizes the patient data for pharmacoepidemiologic analyses. For each comparison, we estimated the unstratified Kaplan-Meier curves, conducted by both propensity score stratified ( $n$  strata = 10) log-rank test and Cox model. After 6 years of follow-up, pioglitazone ( $P = 0.005$ , hazard ratio (HR) = 0.916, 95% confidence interval [CI] 0.861–0.974), febuxostat (HR = 0.815, 95% CI 0.710–0.936,  $P = 0.004$ ), and atenolol (HR = 0.949, 95% CI 0.923–0.976,  $P = 2.8 \times 10^{-4}$ ) are associated with a reduced risk of AD compared with matched control populations (Figs. 5 and 6).

Several clinical trials have been conducted with pioglitazone to treat AD. A phase II study (NCT00982202) showed no statistically significant difference between controls and patients with mild to moderate AD [80]. However, another study showed that pioglitazone was associated with cognitive and functional improvement, as



**Fig. 4** Risk gene-informed drug repurposing for Alzheimer's disease (AD). **a** A Sankey diagram illustrates a global view of 25 repurposable drug candidates with published evidence for AD. These drugs are linked to their physical binding targets or neighborhood proteins derived from network-predicted AD risk genes. **b** Network proximity analysis measures the network distance between disease module and drug targets in the human interactome. A subnetwork indicates the molecular mechanism of pioglitazone implicated in AD, which targets six physical binding proteins of which neighborhoods are 12 predicted AD risk genes. **c** Drugs are grouped by their first-level Anatomical Therapeutic Classification (ATC) codes. The drugs with known anti-AD clinical status, in vitro and in vivo mouse model published data are given. Pioglitazone and febuxostat with anti-AD clinical evidence are highlighted

well stabilization of AD in 42 diabetic patients [81]. Many of these studies were conducted in populations without biological confirmation of AD by biomarkers, and in some cases (e.g., the TOMMOW study; NCT01931566), the dose of pioglitazone was substantially lower than that used in clinical practice for the treatment of diabetes. The available clinical trial data do not exclude a beneficial effect of pioglitazone on AD. Except for the TOMMOW study that was conducted in cognitively normal at-risk individuals, other trials examined symptomatic patients that addressed a question different from the risk-reduction interrogation we prosecuted. For these reasons, we chose pioglitazone to conduct new comparison analysis to reduce unobserved bias.

In pharmacoepidemiologic studies, a comparator drug sharing similar indications with the investigational drug is usually selected as a “control drug” [82]. This approach is able to reduce unobserved bias, as the comparator drug

and the investigational drug are likely to target the same population. Since both of pioglitazone and glipizide are treated for T2D, we therefore selected glipizide as a comparator drug. We next conducted new comparison analyses between pioglitazone and glipizide (an anti-T2D drug,  $n = 191,656$ ) to evaluate the predicted association based on individual-level longitudinal patient data and a novel user active comparator design (“Material and methods”). New comparison analyses confirm that pioglitazone is associated with a reduced risk of AD in comparison to glipizide (HR = 0.921, 95% CI 0.862–0.984,  $P = 0.0159$ , Fig. 5). Thus, two independent comparison analyses support our network-based prediction for pioglitazone.

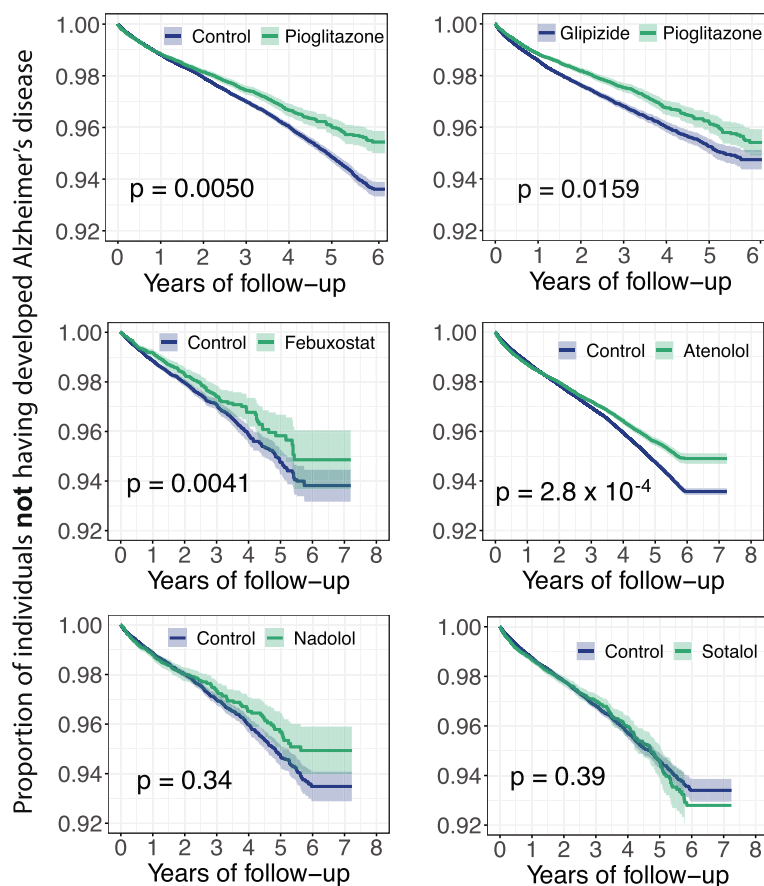
**In vitro observation of pioglitazone’s mechanism-of-action in AD**

Figures 5 and 6 reveal that pioglitazone significantly reduces risk of AD in longitudinal patient-based data.

**Table 2** Statistics of patient data used for pharmacoepidemiologic analysis

Group	Sample size	# of AD	Female (%)	Mean Age (SD)	Geographical location (%)					CAD (%)	T2D (%)	HTN (%)
					Northeast	North Central	South	West	NA			
Pioglitazone (repurposing drug), glipizide (comparative drug) and matched control												
Pioglitazone	101,650	1244	38.5	72.8 (6.9)	17.7	28.4	35.6	17.3	0.9	11.3	51.2	34.2
Glipizide	191,656	3048	44.8	74.2 (7.1)	21.5	27.6	31.4	18.7	0.8	17.1	59.4	43.0
Control	402,488	5230	38.3	74.5 (7.2)	24.2	31.4	30.5	13	1	13.5	44.6	41.1
Febuxostat (repurposing drug), and matched control												
Febuxostat	24,218	243	37.2	75.1 (7.9)	23.5	23.8	37.7	14.4	0.6	28.0	35.3	58.6
Control	95,192	1168	37.0	75.7 (7.4)	24.7	31.6	30.3	12.7	0.8	26.7	34.1	58.1
Atenolol (repurposing drug), and matched control												
Atenolol	366,277	58	58	74.5 (8.0)	24	26	29	21	0.7	11	14	34
Control	1,449,815	58	58	75.2 (7.5)	24	30	30	14	1.1	12	16	38
Nadolol (repurposing drug), and matched control												
Nadolol	19,253	60	60	73.7 (7.8)	29	25	35	10	0.8	11	18	34
Control	76,136	60	60	75.2 (7.5)	24	30	30	14	1.1	12	19	40
Sotalol (repurposing drug), and matched control												
Sotalol	43,819	46	46	76.4 (7.7)	19	29	38	14	0.5	35	22	50
Control	172,375	46	46	76.1 (7.5)	25	31	30	13	0.8	30	22	50

We estimated the unstratified Kaplan-Meier curves, conducted propensity score stratified (*n* strata = 10) log-rank test and Cox model. CAD coronary artery disease, T2D type 2 diabetes, HTN hypertension, SD standard deviation



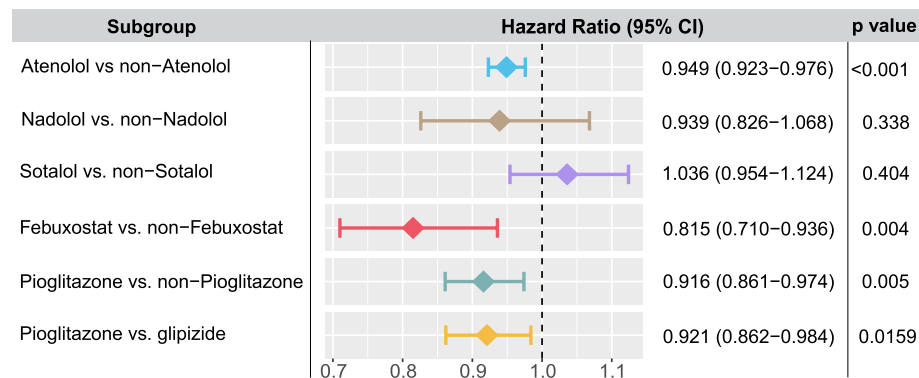
**Fig. 5** Longitudinal analyses reveal that pioglitazone reduces incidence of Alzheimer's disease in patient data. Six comparison analyses were conducted including (i) pioglitazone ( $n = 101,650$ ) vs. matched control population ( $n = 402,184$ ); (ii) pioglitazone vs. glipizide (a diabetes drug,  $n = 191,656$ ); (iii) febuxostat ( $n = 24,218$ ) vs. control ( $n = 95,192$ ); (iv) atenolol ( $n = 366,277$ ) vs. control ( $n = 1,449,815$ ); (v) nadolol ( $n = 19,253$ ) vs. control ( $n = 76,136$ ); and (vi) sotalol ( $n = 43,819$ ) vs. control ( $n = 172,375$ ). First, for each comparison, we estimated the propensity score by using the variables described in Table 2. Then, we estimated the unstratified Kaplan-Meier curves, conducted propensity score stratified ( $n$  strata = 10) log-rank test and Cox model. Using propensity score stratified survival analyses, non-exposures were matched to the exposures (ratio 4:1) by adjusting the initiation time of drug, enrollment history, age and gender, and disease comorbidities (hypertension, type 2 diabetes and coronary artery disease)

To further investigate its mechanism-of-action in AD, we performed a network analysis through integration of drug targets and ARGs into the brain-specific PPI network (see “Material and methods”). Network analysis shows that pioglitazone potentially targets two tauopathy-related proteins (GSK3 $\beta$  and CDK5) in AD (Fig. 7a). RNA sequencing data from the GTEx database (GTEx Consortium, 2015) suggests that GSK3 $\beta$  and CDK5 are highly expressed in brain tissue. Accumulating studies suggested that inhibition of GSK3 $\beta$  and CDK5 activity is a potential therapeutic strategy for AD [83].

We next examined pioglitazone's mechanism-of-action using human microglia HMC3 cells. First, to assess the potential cell cytotoxicity, HMC3 cells were treated with pioglitazone at various concentrations

(0.03  $\mu$ M to 100  $\mu$ M) for 48 h, and cell viability was determined by MTT method. As presented in Fig. 7b, pioglitazone at 0.03  $\mu$ M to 10  $\mu$ M did not affect cell viability, revealing low toxicity in human cells. Thus, these optimized concentrations of pioglitazone ( $\leq 10$   $\mu$ M) were used in subsequent experiments. As shown in Fig. 7c, phosphorylation of GSK3 $\beta$  and CDK5 were significantly increased after LPS treatment (1  $\mu$ g/mL for 30 min) in HMC3 cells. Pre-treating with pioglitazone significantly reduced phosphorylated GSK3 $\beta$  and CDK5 in a dose-dependent manner (Fig. 7d,e and S4). Altogether, these data suggest that pioglitazone may offer potential benefits for patients with AD by reducing activation of GSK3 $\beta$  and CDK5. However, further mechanistic observations using patient-derived





**Fig. 6** Hazard ratios and 95% confidence interval (CI) for six cohort studies. Six cohorts include the following: (i) pioglitazone ( $n = 101,650$ ) vs. matched control population ( $n = 402,184$ ), (ii) pioglitazone vs. glipizide (a diabetes drug,  $n = 191,656$ ); (iii) febuxostat ( $n = 24,218$ ) vs. control ( $n = 95,192$ ), (iv) atenolol ( $n = 366,277$ ) vs. control ( $n = 1,449,815$ ); (v) nadolol ( $n = 19,253$ ) vs. control ( $n = 76,136$ ); and (vi) sotalol ( $n = 43,819$ ) vs. control ( $n = 172,375$ ). For each comparison, we estimated the propensity score for confounding factor (Table 2) adjustment as described in “Material and methods”

microglia cells or disease-relevant cell lines are warranted.

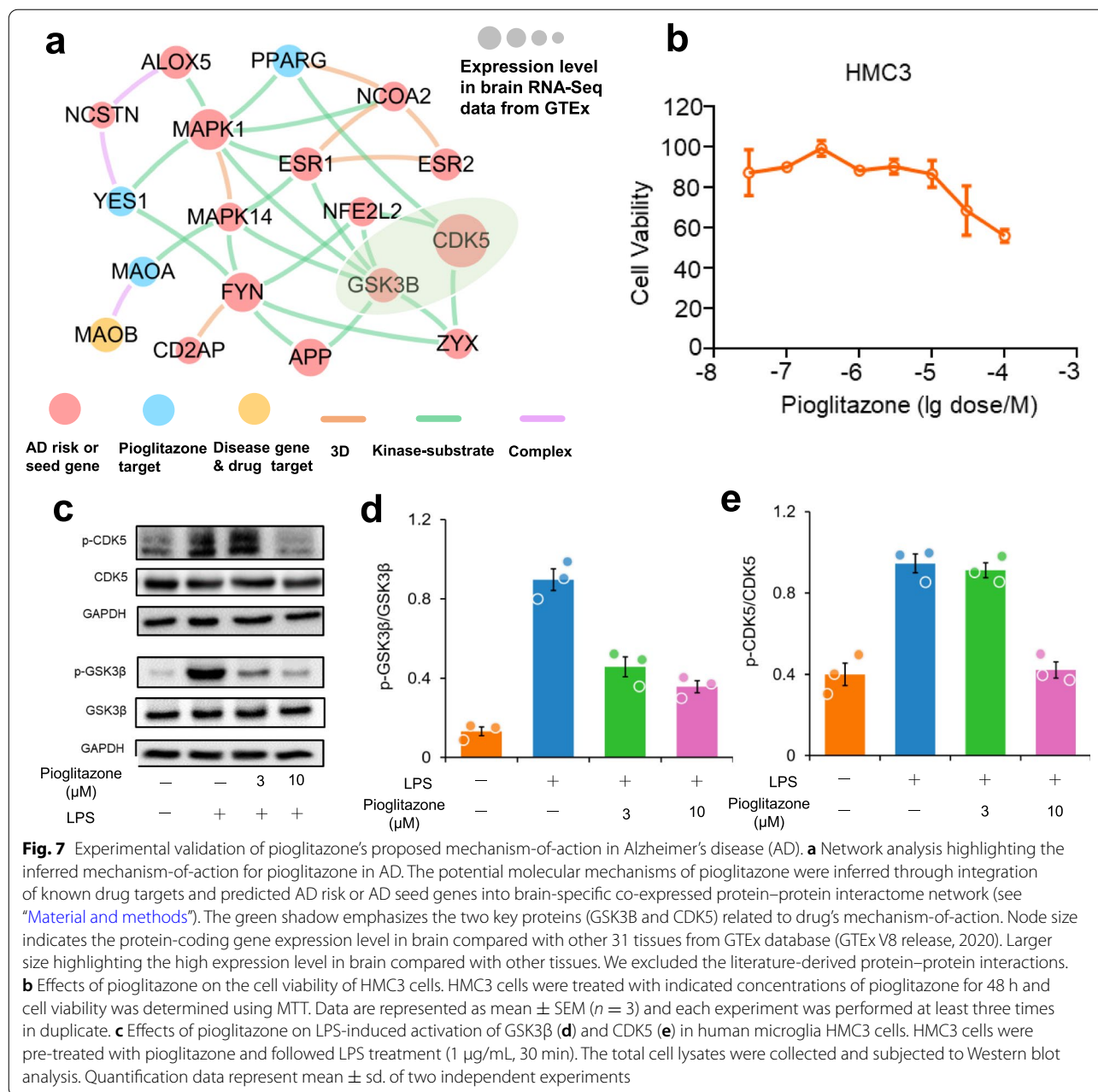
## Discussion

AD risk involves a complex polygenic and pleiotropic genetic architecture [1]. The AD genetics altered the molecular interactions of cellular pathways, which are represented through the organized structure of molecular networks (i.e., gene regulatory networks) or gene co-expression modules [84, 85]. Traditional reductionist paradigms overlook the inherent complexity of human disease and often led to treatments that are inadequate or have adverse effects [86]. Understanding AD from the point-of-view of how cellular systems and molecular interactome perturbations underlie the disease is the essence of network medicine [8]. Based on this hypothesis, we have proposed an artificial intelligence framework for AD drug repurposing, which integrates genetic findings, multi-omics data, drug-target networks, and the human protein–protein interactome, along with large-scale population-based validation and *in vitro* mechanistic observations in human microglia cells (see “Material and methods”).

In total, we identified 103 ARGs by utilizing our recently developed Bayesian model selection method [17]. Functional genomics enrichment analysis shows that ARGs harbor more gene regulatory elements in the human genome. Both transcriptomics and proteomics data analyses imply that ARGs are more likely to be differentially expressed in human AD brain and multiple AD transgenic mouse models. The marginal difference of fold change (Fig. 2) may be explained by large number of cell subpopulations and low abundance of RNA expression at single-cell/nucleus levels during differential

expression analysis. These comprehensive observations suggest that ARGs potentially capture pathobiological pathways of AD (Fig. 3). Importantly, drug-target network analysis shows a 4-fold higher druggability compared to the known drug targets in the human genome. A previous study showed that few products (proteins) of GWAS-derived closest genes could be applied for therapeutic discovery [71].

To compare the performance between the nearest genes to the risk loci using traditional approach and ARGs derived from our Bayesian model, we assembled 108 nearest genes (Table S1) to the GWAS loci data we used (see Supplementary Methods). Unlike to 103 ARGs (Fig. 2a), 108 GWAS-derived closest genes were randomly distributed in the human interactome network ( $7/96$ ,  $P = 0.217$ , permutation test, Table S12). In addition, among 25 candidate drugs (with known AD evidence) identified by ARGs, only 2 drugs (hydralazine and deferoxamine) exerted significant network proximity score with AD (Table S11) based on the 108 nearest genes. None of three positive drugs (pioglitazone [ $z = -1.64$ ], febuxostat [ $z = -1.6$ ], and atenolol [ $z = -1.16$ ]) in Fig. 5 show significant network proximity score with the 108 nearest genes. Altogether, these observations implied poor performance of GWAS-derived closest genes as candidate targets for therapeutic discovery, consistent with previous studies [17, 87]. Several factors may account for this. First, the reported significant loci occupy only a small proportion of heritability and provide limited information about underlying AD biology [88]. Second, many genome-wide significant loci lie in noncoding regions, and genes closest to index SNPs may not represent causal genes of AD [89]. Thus, systematic identification of likely causal genes from GWAS findings



using *in silico* multi-omics approaches is a crucial step for understanding AD pathobiology and offers potential candidate targets for new therapeutic development as presented in this study.

Network-based drug repurposing from ARG findings prioritize 4 repurposable drug candidates for AD, including pioglitazone (NCT02913664), carvedilol (NCT01354444), febuxostat, and fluticasone. Carvedilol, an FDA-approved drug for hypertension that blocks the beta adrenergic receptor, significantly attenuates brain oligomeric β-amyloid level and cognitive deterioration in

two independent AD mice models [90]. Fluticasone is an approved glucocorticoid receptor agonist for treatment of asthma, and recent studies showed that long-term use of fluticasone reduces incidence of developing AD [18, 77]. A propensity-matched analysis has suggested that a daily dose of 40 mg febuxostat is associated with reduced likelihood of dementia in older adults [79]. Nevertheless, whether febuxostat reduces the risk of AD dementia remains unknown. Combining network-based prediction and patient data observation, we found that febuxostat is significantly associated with a decreased risk of AD

(Figs. 5 and 6). In addition, epidemiological studies have shown that hypertension is a risk factor for AD-related dementia; yet, there is some dispute as to whether antihypertensive drugs reduce the risk of AD [91, 92]. Among three adrenergic beta blocker-based antihypertensive drugs (atenolol, nadolol, and sotalol), atenolol is associated with reduced risk of AD, while nadolol and sotalol are not (Figs. 5 and 6). Since there are lack of strong pre-clinical or clinical evidence to support the relationship between atenolol and AD, we excluded atenolol in our follow-up studies. However, future studies to confirm potential beneficial effects of antihypertensive drugs in reducing AD risk are needed.

Pioglitazone, a U.S. FDA-approved anti-T2D drug, was reported to restore energy metabolism and reduce A $\beta$  levels in the brain of APP/PS1 mice [93]. A previous clinical study has shown that pioglitazone improves cognition and regional cerebral blood flow in patients with mild AD accompanied with T2D [81]. In this study, by combining network-based prediction and population-based validation, we found that pioglitazone potentially reduced risk of AD in large-scale patient database (Figs. 5 and 6). Under active drug user design framework [15], we chose a comparator drug having the similar indication with the target drug pioglitazone. As pioglitazone was approved for anti-diabetes, and we therefore chose glipizide as a comparator drug. New active drug user design analysis further support that pioglitazone is associated with a reduced risk of AD in comparison to glipizide. In addition, *in vitro* mechanistic observations (Fig. 7) reveal that pioglitazone significantly downregulates expression of CDK5 and GSK3 $\beta$  in human microglia cells, mechanistically supporting network and population-based findings. However, a phase II study (NCT00982202) shows no statistically significant differences between controls and patients with mild to moderate AD for pioglitazone [80]. One possible explanation is that pioglitazone reduces risk of AD only in patients with pre-existing diabetes or that pioglitazone may have its effects before symptoms occur but not in more advanced patients. Thus, our findings suggest that larger clinical trials and additional mechanistic studies may be necessary to clarify pioglitazone's action in AD prevention in both a broad population and a well-defined subpopulation.

Our network methodology presented here has several strengths. First, it contributes to identification of high-confidence likely causal genes, followed by multi-omics data validation, network-based drug repurposing investigation, large-scale patient data analysis, and *in vitro* mechanistic observation in human microglial cells. This work illustrates translation of GWAS findings to

pathobiology and therapeutic development in AD. Second, our proposed network proximity approach outperform other network approaches, such as weighted gene co-expression network analysis (WGCNA) in which it infers gene co-expression networks from gene/protein expression profiles using network community detection theory [94]. Multiple studies have demonstrated high false positive rate of the gene/protein co-expression networks compared to the physical protein-protein interaction network [95, 96]. Third, the large patient-level longitudinal data ensures that our analyses integrate real-world patient evidence to test the drug's efficiency in AD risk reduction.

### Limitations

Potential weaknesses of this work should be acknowledged. First, as genetic variants from GWAS that influence human disease traits are far from complete, a relative loose threshold ( $1 \times 10^{-5}$ ) rather than genome-wide significant threshold ( $5 \times 10^{-8}$ ) is adopted, which may affect the accuracy of identification of ARGs. Second, we only integrated SNPs associated with AD from large-scale GWAS studies conducted between 2007 and 2019. Since several recent GWAS studies have been conducted [97, 98], we may identify new AD-associated risk genes via integration of the latest novel GWAS loci for AD in the future. Third, incompleteness of human protein-protein interactome network data and potential literature bias may influence performance of our methodology as discussed in a recent study [99]. Since the likely causal genes were predicted based on SNPs identified from GWAS that are primarily centered on the variants in the noncoding regions, some AD genes or proteins harboring protein-coding variants may not be covered in this study. Integration of large-scale whole-genome/exome sequencing from the Alzheimer's Disease Sequencing Project (ADSP; <https://www.niagads.org/adsp/content/home>) may offer novel risk genes for AD. Fourth, detailed clinical information is missing for health insurance claims data regardless of high-dimensional covariate adjustment. This limits our ability to test the effects of pioglitazone on subpopulation of AD patients such as those with mild AD. Longitudinal analyses were conducted in populations without biological confirmation of AD by biomarkers, such as lack of cerebrospinal fluid information (i.e., levels of A $\beta$  and Tau), which may affect the results of pharmacoepidemiologic analyses. In addition, although our dataset contains a geographically diverse population of commercially insured Americans seniors, the results are not representative of individuals who are not commercially insured or uninsured. The phenotyping algorithms may not capture all AD cases. Thus, this approach may need to be re-applied on a regular

iterative basis as datasets are expanded, in order to offer maximum utility. Furthermore, clinical data heterogeneity from real-world patients and other confounding factors may lead to potential false positive rate in population-based drug outcome analysis. Finally, all novel ARGs need to be validated experimentally (including both in vitro and in vivo) and clinical benefits of drugs must be tested in AD randomized clinical trials in the future.

## Conclusion

In summary, this study presents a network-based artificial intelligence methodology to translate GWAS findings to emerging therapeutic discovery by incorporating multi-omics, drug-target network, and the human protein–protein interactome, along with large-scale population-based and in vitro mechanistic observation. This study shows the strong proof-of-concept application of high-confidence risk gene identification from human genetic and multi-omics findings to identifying treatments that can be repurposed for AD and has identified pioglitazone as a potential new treatment for AD using artificial intelligence approaches. In this way, we can minimize the translational gap between genetic findings and clinical outcomes, which is a significant problem in current AD therapeutic development. From a translational perspective, if broadly applied, the artificial intelligence-based tools developed here could help develop novel efficacious treatment strategies for other human complex diseases.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13195-021-00951-z>.

Additional file 1.

## Acknowledgements

We acknowledge the team of the AD Knowledge portal to make genetic and multi-omics data of Alzheimer's disease available for this study.

## Authors' contributions

F.C. designed and conceived the study. J.F., P.Z., and Q.W. performed all experiments and data analysis. Y.Z., Y.H., J.X., R.C., C.W.C., B.Z., B.L., S.J.L., A.A.P., J.B.L., L.L., and J.C., interpreted the data analysis. B.L., Q.W., and R.C. contributed to the inference of AD risk genes from GWAS data. F.C., J.F., P.Z., and Q.W. drafted the manuscript. F.C., J.F., P.Z., A.A.P., and J.C. critically revised the manuscript. All authors read and gave final approval of the manuscript.

## Funding

This work was supported by the National Institute of Aging (NIA) of the National Institutes of Health (NIH) under Award Number R01AG066707, U01AG073323, and 3R01AG066707-01S1 to F.C. This work was supported in part by NIH Research Grant 3R01AG066707-02S1 (F.C.) funded by the Office of Data Science Strategy (ODSS). This work was supported in part by the NIA under Award Number R56AG063870 (F.C. and J.B.L.), R35AG071476 (J.C.), R01AG069900 (B.L.), the Translational Therapeutics Core of the Cleveland Alzheimer's Disease Research Center (P30AG072959 to F.C., J.B.L., A.A.P., and J.C.), NIH-NINDS U01NS093334 (J.C.), NIH-NIGMS P20GM109025 (J.C.), and

NIH-NHGRI U01HG009086 (Q.W., R.C., and B.L.). This work was supported in part by Brockman Foundation (A.A.P.); AHA/Allen Initiative, Grant/Award Number: 19PABH134580006 (A.A.P.).

## Availability of data and materials

Supplementary Material and Methods, 4 Supplementary Figures and 11 Supplementary Tables are freely available at <https://github.com/ChengF-Lab/alzRG>. In addition, all multi-omics data, human protein–protein interactome network, and drug-target network used this study are freely available at <https://alzgps.lerner.ccf.org> as well.

## Declarations

### Ethics approval and consent to participate

Not applicable

### Consent for publication

Not applicable

### Competing interests

Dr. Cummings has provided consultation to Acadia, Actinogen, Alkahest, Alzheon, Annovis, Avanir, Axsome, Biogen, BioXcel, Cassava, Cerecin, Cerevel, Cortexyme, Cytox, EIP Pharma, Eisai, Foresight, GemVax, Genentech, Green Valley, Grifols, Karuna, Merck, Novo Nordisk, Otsuka, Resverlogix, Roche, Samumed, Samus, Signant Health, Suven, Third Rock, and United Neuroscience pharmaceutical and assessment companies. Dr. Cummings has stock options in ADAMAS, AnnovisBio, MedAvante, and BiOasis. Dr. Leverenz has received consulting fees from Vaxxinity, grant support from GE Healthcare and serves on a Data Safety Monitoring Board for Eisai. The other authors declare that they have no competing interests.

### Author details

<sup>1</sup>Genomic Medicine Institute, Lerner Research Institute, Cleveland Clinic, Cleveland, OH 44195, USA. <sup>2</sup>Department of Biostatistics and Health Data Science, School of Medicine, Indiana University, Indianapolis, IN 46202, USA. <sup>3</sup>Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, TN 37212, USA. <sup>4</sup>Vanderbilt Genetics Institute, Vanderbilt University Medical Center, Nashville, TN 37212, USA. <sup>5</sup>Department of Biomedical Informatics, College of Medicine, Ohio State University, Columbus, OH 43210, USA. <sup>6</sup>Department of Pediatrics, Case Western Reserve University, Cleveland, Ohio 44106, USA. <sup>7</sup>Department of Molecular Medicine, Cleveland Clinic Lerner College of Medicine, Case Western Reserve University, Cleveland, OH 44195, USA. <sup>8</sup>Lou Ruvo Center for Brain Health, Neurological Institute, Cleveland Clinic, Cleveland, OH 44195, USA. <sup>9</sup>Harrington Discovery Institute, University Hospitals Cleveland Medical Center, Cleveland, OH 44106, USA. <sup>10</sup>Department of Psychiatry, Case Western Reserve University, Cleveland, OH 44106, USA. <sup>11</sup>Geriatric Psychiatry, GRECC, Louis Stokes Cleveland VA Medical Center, Cleveland, OH 44106, USA. <sup>12</sup>Institute for Transformative Molecular Medicine, School of Medicine, Case Western Reserve University, Cleveland, OH 44106, USA. <sup>13</sup>Department of Neuroscience, Case Western Reserve University, School of Medicine, Cleveland, OH 44106, USA. <sup>14</sup>Chambers-Grundy Center for Transformative Neuroscience, Department of Brain Health, School of Integrated Health Sciences, University of Nevada Las Vegas, Las Vegas, NV 89154, USA. <sup>15</sup>Case Comprehensive Cancer Center, Case Western Reserve University School of Medicine, Cleveland, Ohio 44106, USA.

Received: 28 May 2021 Accepted: 16 December 2021

Published online: 10 January 2022

## References

- Long JM, Holtzman DM. Alzheimer disease: an update on pathobiology and treatment strategies. *Cell*. 2019;179(2):312–39. <https://doi.org/10.1016/j.cell.2019.09.001> pmid: 31564456.
- Masters CL, Bateman R, Blennow K, Rowe CC, Sperling RA, Cummings JL. Alzheimer's disease. *Nat Rev Dis Primers*. 2015;1:15056. <https://doi.org/10.1038/nrdp.2015.56> pmid: 27188934.
- Nichols E, Szoek CEI, Vollset SE, Abbasi N, Abd-Allah F, Abdela J. Global, regional, and national burden of Alzheimer's disease and other



- dementias, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol.* 2019;18(1):88-106. [https://doi.org/10.1016/s1474-4422\(18\)30403-4](https://doi.org/10.1016/s1474-4422(18)30403-4) pmid: 30497964.
4. Jung YJ, Kim YH, Bhalla M, Lee SB. Genomics: new light on Alzheimer's disease research. *Int J Mol Sci.* 2018;19(12):3771. <https://doi.org/10.3390/ijms19123771> pmid: 30486438.
  5. Cuyvers E, Sleegers K. Genetic variations underlying Alzheimer's disease: evidence from genome-wide association studies and beyond. *Lancet Neurol.* 2016;15(8):857-68. [https://doi.org/10.1016/s1474-4422\(16\)00127-7](https://doi.org/10.1016/s1474-4422(16)00127-7) pmid: 27302364.
  6. Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet.* 2013;45(12):1452-8. <https://doi.org/10.1038/ng.2802> pmid: 24162737.
  7. Jansen IE, Savage JE, Watanabe K, Bryois J, Williams DM, Steinberg S. Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nat Genet.* 2019;51(3):404-13. <https://doi.org/10.1038/s41588-018-0311-9> pmid: 30617256.
  8. Fang J, Pieper AA, Nussinov R, Lee G, Bekris L, Leverenz JB. Harnessing endophenotypes and using network medicine in Alzheimer's drug repurposing. *Med Res Rev.* 2020;40(6):2386-426. <https://doi.org/10.1002/med.21709> pmid: 32656864.
  9. Rabinovici GD. Controversy and progress in Alzheimer's disease - FDA approval of aducanumab. *N Engl J Med.* 2021;385(9):771-4. <https://doi.org/10.1056/NEJMp2111320>
  10. Kodamullil AT, Zekri F, Sood M, Hengerer B, Canard L, McHale D, et al. Trial watch: Tracing investment in drug development for Alzheimer disease. *Nat Rev Drug Discovery.* 2017;16(12):819. <https://doi.org/10.1038/nrd.2017.169> pmid: 29056749.
  11. Alteri E, Guizzaro L. Be open about drug failures to speed up research. *Nature.* 2018;563(7731):317-319. <https://doi.org/10.1038/d41586-018-07352-7>.
  12. Cummings JL, Morstorf T, Zhong K. Alzheimer's disease drug-development pipeline: few candidates, frequent failures. *Alzheimers Res Ther.* 2014;6(4):37. <https://doi.org/10.1186/alzrt269> pmid: 25024750.
  13. Nelson MR, Tipney H, Painter JL, Shen J, Nicoletti P, Shen Y, et al. The support of human genetic evidence for approved drug indications. *Nat Genet.* 2015;47(8):856-860. <https://doi.org/10.1038/ng.3314> pmid: 26121088.
  14. Cook D, Brown D, Alexander R, March R, Morgan P, Satterthwaite G, et al. Lessons learned from the fate of AstraZeneca's drug pipeline: a five-dimensional framework. *Nat Rev Drug Discovery.* 2014;13(6):419-431. <https://doi.org/10.1038/nrd4309> pmid: 24833294.
  15. Cheng F, Desai RJ, Handy DE, Wang R, Schneeweiss S, Barabási AL, et al. Network-based approach to prediction and population-based validation of in silico drug repurposing. *Nat Commun.* 2018;9(1):2691. <https://doi.org/10.1038/s41467-018-05116-5> pmid: 30002366.
  16. Cheng F, Lu W, Liu C, Fang J, Hou Y, Handy DE, et al. A genome-wide positioning systems network algorithm for in silico drug repurposing. *Nat Commun.* 2019;10(1):3476. <https://doi.org/10.1038/s41467-019-10744-6> pmid: 31375661.
  17. Wang Q, Chen R, Cheng F, Wei Q, Ji Y, Yang H, et al. A Bayesian framework that integrates multi-omics data and gene networks predicts risk genes from schizophrenia GWAS data. *Nat Neurosci.* 2019;22(5):691-699. <https://doi.org/10.1038/s41593-019-0382-7> pmid: 30988527.
  18. Xu J, Zhang P, Huang Y, Zhou Y, Hou Y, Bekris L, et al. Multimodal single-cell/nucleus RNA sequencing data analysis uncovers molecular networks between disease-associated microglia and astrocytes with implications for drug repurposing in Alzheimer's disease. *Genome Res.* 2021;31(10):1900-1912. <https://doi.org/10.1101/gr.272484.120> pmid: 33627474.
  19. Zhou Y, Fang J, Bekris LM, Kim YH, Pieper AA, Leverenz JB, et al. AlzGPS: a genome-wide positioning systems platform to catalyze multi-omics for Alzheimer's drug discovery. *Alzheimers Res Ther.* 2021;13(1):24. <https://doi.org/10.1186/s13195-020-00760-w> pmid: 33441136.
  20. Boyle EA, Li Yi, Pritchard JK. An expanded view of complex traits: from polygenic to omnigenic. *Cell.* 2017;169(7):1177-1186. <https://doi.org/10.1016/j.cell.2017.05.038> pmid: 28622505.
  21. Mifsud B, Tavares-Cadete F, Young AN, Sugar R. Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. *Nat Genet.* 2015;47(6):598-606. <https://doi.org/10.1038/ng.3286> pmid: 25938943.
  22. Andersson R, Gebhard C, Miguel-Escalada I, Hoof I, Bornholdt J, Boyd M, et al. An atlas of active enhancers across human cell types and tissues. *Nature.* 2014;507(7493):455-461. <https://doi.org/10.1038/nature12787> pmid: 24670763.
  23. Koscielny G, An P, Carvalho-Silva D, Cham JA, Fumis L, Gasparyan R, et al. Open Targets: a platform for therapeutic target identification and validation. *Nucleic Acids Res.* 2017;45(D1):D985-Dd94. <https://doi.org/10.1093/nar/gkw1055> pmid: 27899665.
  24. Keren-Shaul H, Spinrad A, Weiner A, Matcovitch-Natan O, Dvir-Szternfeld R, Ulland TK, et al. A unique microglia type associated with restricting development of Alzheimer's disease. *Cell.* 2017;169(7). <https://doi.org/10.1016/j.cell.2017.05.018> pmid: 28602351.
  25. Zhou Y, Song WM, Andhey PS, Swain A, Levy T, Miller KR, et al. Human and mouse single-nucleus transcriptomics reveal TREM2-dependent and TREM2-independent cellular responses in Alzheimer's disease. *Nat Med.* 2020;26(1):131-142. <https://doi.org/10.1038/s41591-019-0695-9> pmid: 31932797.
  26. Grubman A, Chew G, Ouyang JF. A single-cell atlas of entorhinal cortex from individuals with Alzheimer's disease reveals cell-type-specific gene expression regulation. *Nat Neurosci.* 2019;22(12):2087-2097. <https://doi.org/10.1038/s41593-019-0539-4> pmid: 31768052.
  27. Leng K, Li E, Eser R. Molecular characterization of selectively vulnerable neurons in Alzheimer's disease. *Nat Neurosci.* 2021;24(2):276-287. <https://doi.org/10.1038/s41593-020-00764-7> pmid: 33432193.
  28. Miller JA, Woltjer RL, Goodenbour JM, Horvath S, Geschwind DH. Genes and pathways underlying regional and cell type changes in Alzheimer's disease. *Genome Med.* 2013;5(5):48. <https://doi.org/10.1186/gm452> pmid: 23705665.
  29. Wang M, Roussos P, McKenzie A, Zhou X, Kajiura Y, Brennand KJ, et al. Integrative network analysis of nineteen brain regions identifies molecular signatures and networks underlying selective regional vulnerability to Alzheimer's disease. *Genome Med.* 2016;8(1):104. <https://doi.org/10.1186/s13073-016-0355-3> pmid: 27799057.
  30. Wang Y, Cella M, Mallinson K, Ulrich JD, Young KL, Robinette ML, et al. TREM2 lipid sensing sustains the microglial response in an Alzheimer's disease model. *Cell.* 2015;160(6):1061-1071. <https://doi.org/10.1016/j.cell.2015.01.049> pmid: 25728668.
  31. Orre M, Kamphuis W, Osborn LM, Jansen AHP, Kooijman L, Bossers K, et al. Isolation of glia from Alzheimer's mice reveals inflammation and dysfunction. *Neurobiol Aging.* 2014;35(12):2746-2760. <https://doi.org/10.1016/j.neurobiolaging.2014.06.004> pmid: 25002035.
  32. Politto VA, Li H, Martini-Stoica H, Wang B, Yang L, Xu Y, et al. Selective clearance of aberrant tau proteins and rescue of neurotoxicity by transcription factor EB. *EMBO Mol Med.* 2014;6(9):1142-1160. <https://doi.org/10.15252/emmm.201303671> pmid: 25069841.
  33. Eppig JT, Smith CL, Blake JA, Ringwald M, Kadin JA, Richardson JE, et al. Mouse Genome Informatics (MG): resources for Mining Mouse Genetic, Genomic, and Biological Data in Support of Primary and Translational Research. *Methods Mol Biol (Clifton, NJ).* 2017;1488:47-73. [https://doi.org/10.1007/978-1-4939-6427-7\\_3](https://doi.org/10.1007/978-1-4939-6427-7_3) pmid: 27933520.
  34. Bouter Y, Kacprowski T, Weissmann R, Dietrich K, Borgers H, Brauñ A, et al. Deciphering the molecular profile of plaques, memory decline and neuron loss in two mouse models for Alzheimer's disease by deep sequencing. *Front Aging Neurosci.* 2014;6:75. <https://doi.org/10.3389/fnagi.2014.00075> pmid: 24795628.
  35. Wang H, Li Y, Ryder JW, Hole JT, Ebert PJ, Airey DC, et al. Genome-wide RNAseq study of the molecular mechanisms underlying microglia activation in response to pathological tau perturbation in the rTg4510 tau transgenic animal model. *Mol Neurodegener.* 2018;13(1):65. <https://doi.org/10.1186/s13024-018-0296-y> pmid: 30558641.
  36. Anders S. Analysing RNA-Seq data with the DESeq package. *Mol Biol.* 2010;43(4):1-17.
  37. Savas JN, Wang YZ, DeNardo LA, Martinez-Bartolome S, McClatchy DB, Hark TJ, et al. Amyloid accumulation drives proteome-wide alterations in mouse models of Alzheimer's disease-like pathology. *Cell Rep.* 2017;21(9):2614-2627. <https://doi.org/10.1016/j.celrep.2017.11.009> pmid: 29186695.
  38. Kim DK, Park J, Han D, Yang J, Kim A, Woo J, et al. Molecular and functional signatures in a novel Alzheimer's disease mouse model assessed by

- quantitative proteomics. *Mol Neurodegener.* 2018;13(1):2. <https://doi.org/10.1186/s13024-017-0234-4>. pmid: 29338754.
39. Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, et al. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res.* 2018;46(D1):D1074-D1d82. <https://doi.org/10.1093/nar/gkx1037>. pmid: 29126136.
  40. Li YH, Yu CY, Li XX, Zhang P, Tang J, Yang Q, et al. Therapeutic target database update 2018: enriched resource for facilitating bench-to-clinic research of targeted therapeutics. *Nucleic Acids Res.* 2018;46(D1):D1121-D11d7. <https://doi.org/10.1093/nar/gkx1076>. pmid: 29140520.
  41. Barbarino JM, Whirl-Carrillo M, Altman RB, Klein TE. PharmGKB: a world-wide resource for pharmacogenomic information. *Wiley Interdiscip Rev Syst Biol Med.* 2018;10(4):e1417. <https://doi.org/10.1002/wsbm.1417>. pmid: 29474005.
  42. Wang C, Zhang F, Jiang S, Siedlak SL, Shen L, Perry G, et al. Estrogen receptor-alpha is localized to neurofibrillary tangles in Alzheimer's disease. *Sci Rep.* 2016;6:20352. <https://doi.org/10.1038/srep20352>. pmid: 26837465.
  43. Conejero-Goldberg C, Hyde TM, Chen S, Dreses-Werringloer U, Herman MM, Kleinman JE, et al. Molecular signatures in post-mortem brain tissue of younger individuals at high risk for Alzheimer's disease as based on APOE genotype. *Mol Psych.* 2011;16(8):836-847. <https://doi.org/10.1038/mp.2010.57>. pmid: 20479757.
  44. Granic I, Dolga AM, Nijholt IM, van Dijk G, Eisel UL. Inflammation and NF-kappaB in Alzheimer's disease and diabetes. *J Alzheimer's Dis.* 2009;16(4):809-821. <https://doi.org/10.3233/jad-2009-0976>. pmid: 19387114.
  45. Ma QL, Yang F, Frautschy SA, Cole GM. PAK in Alzheimer disease, Huntington disease and X-linked mental retardation. *Cell Logist.* 2012;2(2):117-125. <https://doi.org/10.4161/cl.21602>. pmid: 23162743.
  46. Zhao J, Cheng F, Jia P, Cox N, Denny JC, Zhao Z. An integrative functional genomics framework for effective identification of novel regulatory variants in genome-phenome studies. *Genome Med.* 2018;10(1):7. <https://doi.org/10.1186/s13073-018-0513-x>. pmid: 29378629.
  47. Won H, de la Torre-Ubieta L, Stein JL, Parikhshak NN, Huang J, Opland CK, et al. Chromosome conformation elucidates regulatory relationships in developing human brain. *Nature.* 2016;538(7626):523-527. <https://doi.org/10.1038/nature19847>. pmid: 27760116.
  48. Hesse R, Hurtado ML, Jackson RJ, Eaton SL, Herrmann AG, Colom-Cadena M, et al. Comparative profiling of the synaptic proteome from Alzheimer's disease patients with focus on the APOE genotype. *Acta Neuropathologica Commun.* 2019;7(1):214. <https://doi.org/10.1186/s40478-019-0847-7>. pmid: 31862015.
  49. Seyfried NT, Dammer EB, Swarup V, Nandakumar D, Duong DM, Yin L, et al. A multi-network approach identifies protein-specific co-expression in asymptomatic and symptomatic Alzheimer's disease. *Cell Syst.* 2017;4(1):60-72.e4. <https://doi.org/10.1016/j.cels.2016.11.006>. pmid: 27989508.
  50. Wingo AP, Dammer EB. Large-scale proteomic analysis of human brain identifies proteins associated with cognitive trajectory in advanced age. *Nature Commun.* 2019;10(1):1619. <https://doi.org/10.1038/s41467-019-09613-z>. pmid: 30962425.
  51. Mathys H, Davila-Velderrain J, Peng Z, Gao F, Mohammadi S, Young JZ, et al. Single-cell transcriptomic analysis of Alzheimer's disease. *Nature.* 2019;570(7761):332-337. <https://doi.org/10.1038/s41586-019-1195-2>. pmid: 31042697.
  52. Zhao L, Ma QL, Calon F, Harris-White ME, Yang F, Lim GP, et al. Role of p21-activated kinase pathway defects in the cognitive deficits of Alzheimer disease. *Nat Neurosci.* 2006;9(2):234-242. <https://doi.org/10.1038/nn1630>. pmid: 16415866.
  53. Borjes C, Arsenault D, Lemire M, Tremblay C, De Koninck Y, Calon F. Transgenic autoinhibition of p21-activated kinase exacerbates synaptic impairments and fronto-dependent behavioral deficits in an animal model of Alzheimer's disease. *Aging.* 2017;9(5):1386-1403. <https://doi.org/10.18632/aging.101239>. pmid: 28522792.
  54. Arsenault D, Dal-Pan A, Tremblay C, Bennett DA, Guitton MJ, De Koninck Y, et al. PAK inactivation impairs social recognition in 3xTg-AD Mice without increasing brain deposition of tau and Abeta. *J Neurosci.* 2013;33(26):10729-10740. <https://doi.org/10.1523/jneurosci.1501-13.2013>. pmid: 23804095.
  55. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, et al. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics (Oxford, England).* 2009;25(8):1091-3. <https://doi.org/10.1093/bioinformatics/btp101>. pmid: 19237447.
  56. Allen M, Kachadoorian M, Carrasquillo MM, Karhade A, Manly L, Burgess JD, et al. Late-onset Alzheimer disease risk variants mark brain regulatory loci. *Neurol Genet.* 2015;1(2):e15. <https://doi.org/10.1212/nxg.000000000000012>. pmid: 27066552.
  57. Karch CM, Goate AM. Alzheimer's disease risk genes and mechanisms of disease pathogenesis. *Biol Psychiat.* 2015;77(1):43-51. <https://doi.org/10.1016/j.biopsych.2014.05.006>. pmid: 24951455.
  58. Hondius DC, van Nierop P, Li KW, Hoozemans JJ, van der Schors RC, van Haastert ES, et al. Profiling the human hippocampal proteome at all pathologic stages of Alzheimer's disease. *Alzheimer's Dement.* 2016;12(6):654-668. <https://doi.org/10.1016/j.jalz.2015.11.002>. pmid: 26772638.
  59. Kunkle BW, Grenier-Boley B, Sims R, Bis JC, Damotte V, Naj AC, et al. Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Abeta, tau, immunity and lipid processing. *Nat Genet.* 2019;51(3):414-430. <https://doi.org/10.1038/s41588-019-0358-2>. pmid: 30820047.
  60. Fjorback AW, Seaman M, Gustafsen C, Mehmedbasic A, Gokool S, Wu C, et al. Retromer binds the FANSHY sorting motif in SorLa to regulate amyloid precursor protein sorting and processing. *J Neurosci.* 2012;32(4):1467-1480. <https://doi.org/10.1523/jneurosci.2272-11.2012>. pmid: 22279231.
  61. Peron R, Vatanabe IP, Manzine PR, Camins A, Cominetti MR. Alpha-secretase ADAM10 regulation: insights into Alzheimer's disease treatment. *Pharmaceuticals (Basel, Switzerland).* 2018;11(1). <https://doi.org/10.3390/ph11010012>. pmid: 29382156.
  62. Farhat SM, Ahmed T. Neuroprotective and neurotoxic implications of alpha7 nicotinic acetylcholine receptor and beta interaction: therapeutic options in Alzheimer's disease. *Cur Drug Targets.* 2017;18(13):1537-1544. <https://doi.org/10.2174/1389450117666161005145143>. pmid: 27719660.
  63. Wang LL, Pan XL, Wang Y, Tang HD, Deng YL, Ren RJ, et al. A single nucleotide polymorphism in LRP2 is associated with susceptibility to Alzheimer's disease in the Chinese population. *Clinica Chimica Acta.* 2011;412(3-4):268-270. <https://doi.org/10.1016/j.cca.2010.10.015>. pmid: 20971101.
  64. Dineley KT, Westerman M, Bui D, Bell K, Ashe KH, Sweatt JD. Beta-amyloid activates the mitogen-activated protein kinase cascade via hippocampal alpha7 nicotinic acetylcholine receptors: In vitro and in vivo mechanisms related to Alzheimer's disease. *J Neurosci.* 2001;21(12):4125-4133. pmid: 11404397.
  65. Dourlen P, Fernandez-Gomez FJ, Dupont C, Grenier-Boley B, Bellenguez C, Obriot H, et al. Functional screening of Alzheimer risk loci identifies PTK2B as an in vivo modulator and early marker of Tau pathology. *Mol Psychiat.* 2017;22(6):874-883. <https://doi.org/10.1038/mp.2016.59>. pmid: 27113998.
  66. Salazar SV, Cox TO, Lee S, Brody AH, Chyung AS, Haas LT, et al. Alzheimer's disease risk factor Pyk2 mediates amyloid-beta-induced synaptic dysfunction and loss. *J Neurosci.* 2019;39(4):758-772. <https://doi.org/10.1523/jneurosci.1873-18.2018>. pmid: 30518596.
  67. Jiang T, Sun Q, Chen S. Oxidative stress: a major pathogenesis and potential therapeutic target of antioxidative agents in Parkinson's disease and Alzheimer's disease. *Pro Neurobiol.* 2016;147:1-19. <https://doi.org/10.1016/j.pneurobio.2016.07.005>. pmid: 27769868.
  68. Shi C, Viccaro K, Lee HG, Shah K. Cdk5-Foxo3 axis: initially neuroprotective, eventually neurodegenerative in Alzheimer's disease models. *J Cell Sci.* 2016;129(9):1815-1830. <https://doi.org/10.1242/jcs.185009>. pmid: 28157684.
  69. Snow WM, Albenis BC. Neuronal gene targets of NF-kappaB and their dysregulation in Alzheimer's disease. *Front Mol Neurosci.* 2016;9:118. <https://doi.org/10.3389/fnmol.2016.00118>. pmid: 27881951.
  70. Okada Y, Wu D, Trynka G, Raj T, Terao C, Ikari K, et al. Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature.* 2014;506(7488):376-381. <https://doi.org/10.1038/nature12873>. pmid: 24390342.
  71. Kwok MK, Lin SL, Schooling CM. Re-thinking Alzheimer's disease therapeutic targets using gene-based tests. *EBioMedicine.* 2018;37:461-470. <https://doi.org/10.1016/j.ebiom.2018.10.001>. pmid: 30314892.

72. Philibin SD, Walentiny DM, Vunck SA, Prus AJ, Meltzer HY, Porter JH. Further characterization of the discriminative stimulus properties of the atypical antipsychotic drug clozapine in C57BL/6 mice: role of 5-HT(2A) serotonergic and alpha (1) adrenergic antagonism. *Psychopharmacology*. 2009;203(2):303-315. <https://doi.org/10.1007/s00213-008-1385-3>. pmid: 18989659.
73. Choi Y, Jeong HJ, Liu QF, Oh ST, Koo BS, Kim Y, et al. Clozapine improves memory impairment and reduces abeta level in the Tg-APPswe/PS1dE9 mouse model of Alzheimer's disease. *Mol Neurobiol*. 2017;54(1):450-460. <https://doi.org/10.1007/s12035-015-9636-x>. pmid: 26742522.
74. Kobayashi K, Yamanaka Y, Iwazaki N, Nakajo I, Hosokawa M, Negishi M, et al. Identification of HMG-CoA reductase inhibitors as activators for human, mouse and rat constitutive androstane receptor. *Drug Metabol Dis*. 2005;33(7):924-929. <https://doi.org/10.1124/dmd.104.002741>. pmid: 15802384.
75. Li Y, Liu Q, Sun J, Wang J, Liu X, Gao J. Mitochondrial protective mechanism of simvastatin protects against amyloid beta peptide-induced injury in SH-SY5Y cells. *Int J Mol Med*. 2018;41(5):2997-3005. <https://doi.org/10.3892/ijmm.2018.3456>. pmid: 29436584.
76. Yamamoto N, Fujii Y, Kasahara R, Tanida M, Ohora K, Ono Y, et al. Simvastatin and atorvastatin facilitates amyloid beta-protein degradation in extracellular spaces by increasing neprilysin secretion from astrocytes through activation of MAPK/Erk1/2 pathways. *Glia*. 2016;64(6):952-962. <https://doi.org/10.1002/glia.12974>. pmid: 26875818.
77. Lehrer S, Rheinstein PH. Alzheimer's disease and intranasal fluticasone propionate in the FDA MedWatch Adverse Events Database. *J Alzheimer's Dis Rep*. 2018;2(1):111-115. <https://doi.org/10.3233/adr-170033>. pmid: 30159547.
78. Galimberti D, Scarpini E. Pioglitazone for the treatment of Alzheimer's disease. *Exp Opin Invest Drugs*. 2017;26(1):97-101. <https://doi.org/10.1080/13543784.2017.1265504>. pmid: 27885860.
79. Singh JA, Cleveland JD. Comparative effectiveness of allopurinol versus febuxostat for preventing incident dementia in older adults: a propensity-matched analysis. *Arthritis Res Ther*. 2018;20(1):167. <https://doi.org/10.1186/s13075-018-1663-3>. pmid: 30075731.
80. Geldmacher DS, Fritsch T, McClendon MJ, Landreth G. A randomized pilot clinical trial of the safety of pioglitazone in treatment of patients with Alzheimer disease. *Arch Neurol*. 2011;68(1):45-50. <https://doi.org/10.1001/archneurol.2010.229>. pmid: 20837824.
81. Sato T, Hanyu H, Hira K, Kanetaka H, Sakurai H, Iwamoto T. Efficacy of PPAR-gamma agonist pioglitazone in mild Alzheimer disease. *Neurobiol Aging*. 2011;32(9):1626-1633. <https://doi.org/10.1016/j.neurobiolaging.2009.10.009>. pmid: 19923038.
82. Desai RJ, Franklin JM. Alternative approaches for confounding adjustment in observational studies using weighting based on the propensity score: a primer for practitioners. *BMJ (Clinical research ed)*. 2019;367:15657. <https://doi.org/10.1136/bmj15657>. pmid: 31645336.
83. Mazanetz MP, Fischer PM. Untangling tau hyperphosphorylation in drug design for neurodegenerative diseases. *Nat Rev Drug Discovery*. 2007;6(6):464-479. <https://doi.org/10.1038/nrd2111>. pmid: 17541419.
84. Swarup V, Hinz FI, Rexach JE, Noguchi KI, Toyoshiba H, Oda A, et al. Identification of evolutionarily conserved gene networks mediating neurodegenerative dementia. *Nat Med*. 2019;25(1):152-164. <https://doi.org/10.1038/s41591-018-0223-3>. pmid: 30510257.
85. Zhang B, Gaiteri C, Bodea LG, Wang Z, McElwee J, Podtelezhnikov AA, et al. Integrated systems approach identifies genetic nodes and networks in late-onset Alzheimer's disease. *Cell*. 2013;153(3):707-720. <https://doi.org/10.1016/j.cell.2013.03.030>. pmid: 23622250.
86. Greene JA, Loscalzo J. Putting the patient back together - social medicine, network medicine, and the limits of reductionism. *N Engl J Med*. 2017;377(25):2493-2499. <https://doi.org/10.1056/NEJMms1706744>. pmid: 29262277.
87. Reay WR, Cairns MJ. Advancing the use of genome-wide association studies for drug repurposing. *Nat Rev Genet*. 2021;22(10):658-671. <https://doi.org/10.1038/s41576-021-00387-z>. pmid: 34302145.
88. Deming Y, Li Z, Kapoor M, Harari O, Del-Aguila JL, Black K, et al. Genome-wide association study identifies four novel loci associated with Alzheimer's endophenotypes and disease modifiers. *Acta Neuropathologica*. 2017;133(5):839-856. <https://doi.org/10.1007/s00401-017-1685-y>. pmid: 28247064.
89. Zhang F, Lupski JR. Non-coding genetic variants in human disease. *Hu Mol Genet*. 2015;24(R1):R102-R110. <https://doi.org/10.1093/hmg/ddv259>. pmid: 26152199.
90. Wang J, Ono K, Dickstein DL, Arrieta-Cruz I, Zhao W, Qian X, et al. Carvedilol as a potential novel agent for the treatment of Alzheimer's disease. *Neurobiol Aging*. 2011;32(12):2321.e1-12. <https://doi.org/10.1016/j.neurobiolaging.2010.05.004>. pmid: 20579773.
91. Ding J, Davis-Plourde KL, Sedaghat S, Tully PJ, Wang W, Phillips C, et al. Antihypertensive medications and risk for incident dementia and Alzheimer's disease: a meta-analysis of individual participant data from prospective cohort studies. *Lancet Neurol*. 2020;19(1):61-70. [https://doi.org/10.1016/s1474-4422\(19\)30393-x](https://doi.org/10.1016/s1474-4422(19)30393-x). pmid: 31706889.
92. Hu Z, Wang L, Ma S, Kirisci L, Feng Z, Xue Y, et al. Synergism of antihypertensives and cholinesterase inhibitors in Alzheimer's disease. *Alzheimer's Dement (N Y)*. 2018;4:542-555. <https://doi.org/10.1016/j.trci.2018.09.001>. pmid: 30386819.
93. Chang KL, Wong LR, Pee HN, Yang S, Ho PC. Reverting metabolic dysfunction in cortex and cerebellum of APP/PS1 mice, a model for Alzheimer's disease by pioglitazone, a peroxisome proliferator-activated receptor gamma (PPARgamma) Agonist. *Mol Neurobiol*. 2019;56(11):7267-7283. <https://doi.org/10.1007/s12035-019-1586-2>. pmid: 31016475.
94. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform*. 2008;9:559. <https://doi.org/10.1186/1471-2105-9-559>. pmid: 19114008.
95. Paci P, Fiscon G, Conte F, Wang RS, Farina L, Loscalzo J. Gene co-expression in the interactome: moving from correlation toward causation via an integrated approach to disease module discovery. *NPJ Syst Biol Applicat*. 2021;7(1):3. <https://doi.org/10.1038/s41540-020-00168-0>. pmid: 33479222.
96. Menche J, Sharma A, Kitsak M, Ghiassian SD, Vidal M, Loscalzo J, et al. Disease networks. Uncovering disease-disease relationships through the incomplete interactome. *Science (New York, NY)*. 2015;347(6224):1257601. <https://doi.org/10.1126/science.1257601>. pmid: 25700523.
97. Wightman DP, Jansen IE, Savage JE. A genome-wide association study with 1,126,563 individuals identifies new risk loci for Alzheimer's disease. *Nat Genet*. 2021;53(9):1276-1282. <https://doi.org/10.1038/s41588-021-00921-z>. pmid: 34493870.
98. Schwartzentruber J, Cooper S, Liu JZ, Barrio-Hernandez I. Genome-wide meta-analysis, fine-mapping and integrative prioritization implicate new Alzheimer's disease risk genes. *Nat Genet*. 2021;53(3):392-402. <https://doi.org/10.1038/s41588-020-00776-w>. pmid: 33589840.
99. Fang J, Zhang P, Zhou Y, Chiang WC, Tan J, Hou Y, Stauffer S, Li L, Pieper AA, Cummings J, Cheng F. Endophenotype-based in-silico network medicine discovery combined with insurance records data mining identifies sildenafil as a candidate drug for Alzheimer's disease. *Nat Aging*. 2021;1:1175-88. <https://doi.org/10.1038/s43587-021-00138-z>.

## Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

