An Integrated Molecular Atlas of Alzheimer's Disease

Maria A. Wörheide[1], Jan Krumsiek[2], Serge Nataf[3], Kwangsik Nho[4], Anna K. Greenwood[5], Tong Wu[1], Kevin Huynh[6], Patrick Weinisch[1], Werner Römisch-Margl[1], Nick Lehner[1], The AMP-AD Consortium[+], The Alzheimer's Disease Neuroimaging Initiative[&], The Alzheimer's Disease Metabolomics Consortium[#], Jan Baumbach[7,8], Peter J. Meikle[6], Andrew J. Saykin[4], P. Murali Doraiswamy[9,10,11], Cornelia van Duijn[12,13], Karsten Suhre[14], Rima Kaddurah-Daouk[9,10,11], Gabi Kastenmüller[1,15,*], Matthias Arnold[1,9,*]

[1] Institute of Computational Biology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany

[2] Institute for Computational Biomedicine, Englander Institute for Precision Medicine, Department of Physiology and Biophysics, Weill Cornell Medicine, New York, NY, USA

[3] Univ Lyon, Hospices Civils de Lyon, Université Claude Bernard Lyon 1, Inserm, Stem Cell and Brain Research Institute U1208, F-69500 Bron, France

[4] Department of Radiology and Imaging Sciences, Center for Computational Biology and Bioinformatics, and the Indiana Alzheimer Disease Center, Indiana University School of Medicine, Indianapolis, IN, USA

[5] Sage Bionetworks, Seattle, WA, USA

[6] Metabolomics Laboratory, Baker Heart and Diabetes Institute, Melbourne, VIC 3004, Australia

[7] Institute of Mathematics and Computer Science, University of Southern Denmark, 5230 Odense, Denmark

[8] Chair of Computational Systems Biology, University of Hamburg, 22607 Hamburg, Germany

[9] Department of Psychiatry and Behavioral Sciences, Duke University, Durham, NC, USA

[10] Duke Institute of Brain Sciences, Duke University, Durham, NC, USA

[11] Department of Medicine, Duke University, Durham, NC, USA

[12] Department of Epidemiology, Erasmus Medical Center, Rotterdam, the Netherlands

[13] Nuffield Department of Population Health, University of Oxford, Oxford, UK

[14] Department of Physiology and Biophysics, Weill Cornell Medicine – Qatar, Education City, Doha, Qatar

[15] German Center for Diabetes Research (DZD), Neuherberg, Germany

[+] The AMP-AD Consortium:

The full list of contributing scientists is available at

https://adknowledgeportal.org/AMPADConsortiumMembers.

[&] The Alzheimer's Disease Neuroimaging Initiative (ADNI):

Data used in preparation of this article were obtained from the ADNI database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at:

http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

[#] The Alzheimer's Disease Metabolomics Consortium (ADMC):

The full list of contributing scientists is available at https://sites.duke.edu/adnimetab/team/.

[*] Corresponding author information:

Matthias Arnold, Ph.D.

Department of Psychiatry and Behavioral Sciences

Duke University Medical Center

Durham, NC, USA

Tel: +1-919-684-2611

E-Mail: matthias.arnold@adatlas.org

Gabi Kastenmüller, Ph.D.

Institute of Computational Biology

Helmholtz Zentrum München

Ingolstädter Landstraße 1

D-85764 Neuherberg, Germany

Tel: +49-89-3187-3578

E-Mail: g.kastenmueller@adatlas.org

ABSTRACT

INTRODUCTION: Embedding single-omics disease associations into the wider context of multi-level molecular changes in Alzheimer's disease (AD) remains one central challenge in AD research.

METHODS: Results from numerous AD-specific omics studies from AMP-AD, NIAGADS, and other initiatives were integrated into a comprehensive network resource and complemented with molecular associations from large-scale population-based studies to provide a global view on AD.

RESULTS: We present the AD Atlas, an online resource (www.adatlas.org) integrating over 20 large studies providing disease-relevant information on 20,353 protein-coding genes, 8,615 proteins, 997 metabolites and 31 AD-related phenotypes. Multiple showcases demonstrate the utility of this resource for contextualization of AD research results and subsequent downstream analyses, such as drug repositioning approaches.

DISCUSSION: By providing a global view on multi-omics results through a user-friendly interface, the AD Atlas enables the formulation of molecular hypotheses and retrieval of clinically relevant insights that can be validated in follow-up analyses or experiments.

KEYWORDS

Alzheimer's disease; multi-omics; network medicine; drug repositioning; online resource; data integration

1 Introduction

Late-onset Alzheimer's disease (AD) is a progressive neurodegenerative disorder, for which there is currently no cure or preventive therapy and only modestly effective symptomatic treatments [1]. The failures of hundreds of trials of disease-modifying therapeutics, including several phase III trials targeting amyloid-beta (Aβ), and availability of only one FDA-approved anti-amyloid compound [2] highlight our incomplete understanding of both the cause of AD and the mechanisms of cognitive failure [3]. AD is a multifactorial disease with a long prodromal period as well as substantial heterogeneity in both risk profiles and clinical/pathological presentation. It is linked to all molecular layers from genetic and epigenetic variation through transcriptional changes to altered abundances of proteins and metabolites, which interact in complex networks [4]. Hence, AD is best viewed as a complex alteration in many molecular readouts, which can be seen as a shift of a multi-molecular network from a "normal" to a perturbed state.

Despite significant advances in the study of AD and related dementias, there are many challenges remaining as recently highlighted in the 2021 National Institutes of Health (NIH) AD research summit. One of the most prominent missing pieces are robust and reliable biomarkers for both diagnosis and therapeutic intervention that are embedded in the context of multi-level molecular changes observed in AD and are evaluated in an open, rigorous and reproducible manner. NIH's Accelerating Medicines Partnership in AD (AMP-AD; https://www.nia.nih.gov/research/amp-ad; [5]) program is working towards this goal through generation and examination of diverse data including multi-omics profiling of different modalities and across relevant tissues, where all data generated through the AMP-AD initiative is rapidly shared through the AD Knowledge Portal (https://adknowledgeportal.org; [6]).

The AD Knowledge Portal's Agora Platform (https://agora.ampadportal.org) provides interactive visualizations designed to support the evaluation of data from RNA-seq, proteomics, and metabolomics studies on the single target level. However, a user-friendly analytical tool that

incorporates single biological entities into their multi-omics context has so far been missing. To this end, networks offer an intuitive framework to integrate and store densely connected biomedical data, making them an attractive data structure for multi-omics integration efforts [7]. Heterogeneous networks, which consist of multiple types of nodes (e.g. metabolites, genes and phenotypes) and edges (e.g. partial correlation of metabolites, gene co-expression), have been particularly useful to describe the complex interplay within and between biological domains [8]. Such network-based multi-omics approaches have the potential to ultimately construct comprehensive and largely bias-free models of AD that can guide the identification and prioritization of potential therapeutic targets and drug repositioning candidates [9,10] as well as inform novel hypotheses that can be tested in follow-up experiments.

We here present the AD Atlas, a network-based data integration resource for investigating AD, its biomarkers, and associated (endo-)phenotypes in a multi-omics context. Using an extended QTL-based integration strategy combined with a composite network approach [11], the AD Atlas integrates data from more than 20 studies. Based on data from knowledge bases and healthy cohorts, we constructed a generalized, disease-independent high-quality framework of intra- (e.g. gene-gene) and inter-omics (e.g. metabolite-gene) relationships. Using large-scale association data of AD – including data from AMP-AD, NIAGADS, and other large studies and consortium efforts – this framework was then transformed into an integrated multi-omics knowledgebase for markers of AD. The resulting comprehensive catalogue of multi-omics relationships, stored using the graph-based database Neo4j, provides disease-relevant information on over 20,000 protein-coding genes, 8,000 proteins and nearly 1,000 metabolites as well as associated genetic variants. Lastly, we have developed a publicly available network- and web-based user interface featuring several data analysis tools (www.adatlas.org) to enable access to these complex data independent of in-house bioinformatics capacities. The AD Atlas allows users to construct, expand and explore context- and tissue-specific molecular subnetworks surrounding either individual or multiple entities of interest.

We demonstrate the utility of this resource to generate disease-relevant insights using a variety of showcases ranging from hypothesis-driven to hypothesis-generating analyses.

2 Methods

Core methods used to build the AD Atlas and user interface are briefly described in the following. For full details, including a comprehensive list of integrated data sets, please refer to the **Supplementary Material**.

2.1 Data integration and storage

The AD Atlas was built using step-wise data integration in which different omics datasets, such as transcriptomics, proteomics and metabolomics, are analysed separately or in specific combinations before integration (**Figure 1A**). In contrast to synchronous integration, where all data is used in one analysis step, step-wise approaches allow the integration of data across many different sources and do not require the data to come from the same set of individuals/samples. We used an extended QTL-based integration strategy paired with a composite network approach, described in detail in our recent review [11]. Briefly, relationships between biological entities (e.g. genetic variants, genes, metabolites) are either taken from public knowledge databases or inferred through statistical analysis (e.g. genome-wide association studies (GWAS) or correlation-based analysis). These individual networks are subsequently merged into a large heterogeneous network by using the overlap of common entities, and links between metabolites and genes are established via overlapping quantitative trait loci (QTLs). The resulting heterogeneous network consist of multiple node types (biological entities) that are connected by different types of edges (inferred associations between entities). To enable efficient data storage and analysis we utilized the native graph database management system Neo4j (https://neo4j.com/).

2.2 Data collection and preprocessing

Biological relationships between omics layers were downloaded from public databases, such as gene-transcript-protein mappings from Ensembl [12] and mappings between single nucleotide polymorphisms (SNPs) and genes from SNiPA [13]. SNPs, genes, transcripts and proteins were stored in the database as individual nodes and relationships were added as edges between them. Furthermore, metabolites (and their corresponding meta-information) measured by metabolomics platforms were also collected and stored as nodes in the database. Large-scale quantitative data from population-based studies were then used to establish data-driven relationships within (e.g. tissue-specific gene co-expression) and across omics (e.g. expression QTLs, protein QTLs, or metabolite QTLs) layers. To identify entities within this network that are relevant to AD, we used large-scale association data for AD yielded in case-control and AD biomarker GWASs, metabolome-wide association studies (MWASs), data on differentially expressed genes and differentially abundant proteins, and brain region-specific gene and protein co-expression. A summary of data types and their respective sources is listed in **Table 1**. Summary statistics of these analyses were either downloaded from publicly available data repositories, supplementary data or calculated in an additional analysis step (see the **Supplementary Material** for more details). Each biological entity (e.g. gene, protein, metabolite) was mapped to a unique identifier, either using the mapping provided in the data source or through manual curation. The unique identifier for SNPs is defined as their rsID, while genes, transcripts and proteins are identified by their respective Ensembl IDs and measured metabolites are identified by their platform-specific ID. Additional identifiers, including biochemical metabolite names, gene symbols and UniProt identifiers, have also been annotated but are not required to be unique. AD-specific (endo-)phenotypes from different studies (GWAS, MWAS) were harmonized through manual curation and are listed in **Supplementary Table 3**.

2.3 Abstraction layer - node mapping and edge filtering

SNPs, transcripts and encoded proteins were projected onto genes using information on genomic location, regulatory elements and quantitative trait loci as provided by Ensembl [12], GTEx [14] and SNiPA [13]. Metabolites that were measured on more than one platform were consolidated using manual mappings between platform specific IDs and information on identified unknown metabolites. Statistical associations between biological entities (edges) were summarized and pre-filtered using study specific significance thresholds, with the exception of genetic associations with AD traits (traitQTLs) and metabolites (mQTLs), where either a genome-wide (p-value $\leq 5x10^{-8}$) or gene-wise cutoff can be applied through our web interface. The gene-wise cutoff is defined as p-value $\leq 0.05/\#SNPs_{geneA}$, where $\#SNPs_{geneA}$ is the number of SNPs that have been annotated to $gene_A$. A comprehensive list of study-specific significance thresholds and a more detailed description of the summary and abstraction process is provided in the **Supplementary Material** and **Supplementary Table 2**.

2.4 Implementation

The integrated multi-omics data is stored as a heterogeneous graph using the graph database management system Neo4j (community v4.2.1). To enable easy access to the data, we implemented a network- and web-based user interface to the AD Atlas. This frontend was built as a ShinyApp using R v3.6.2 and is deployed using ShinyProxy (v2.3.1_amd64.deb). Communication between Neo4j and R is established through the official Neo4j python driver *neo4j* (https://github.com/neo4j/neo4j-python-driver) and the R-to-python interface package *reticulate* (v1.18). Interactive networks are visualized using *VisNetwork* (v2.0.9). Enrichment analysis of the generated subnetworks can be performed using the R package *enrichR* (v3.0) [15], which enables the analysis of a variety of different gene sets, including drug perturbation signatures and biochemical pathways, via the Enrichr webservice (https://maayanlab.cloud/Enrichr/) or the R package *gprofiler2* (v0.2.0) [16], which provides an interface to the gene list functional profiling toolset, g:Profiler

(https://biit.cs.ut.ee/gprofiler/gost). Classical GO term enrichment analysis is provided via the R package *topGO* (v2.38.1) [17]. Metabolic pathway enrichment analysis for the metabolites in the constructed subnetworks is performed using platform-specific annotations of metabolites into classes and super-/sub-pathways where this information is available. Contingency tables are calculated and then tested using Fisher's exact test (function *fisher.test* with alternative = 'greater' from R package *stats*). Multiple testing correction is performed using *p.adjust* (method = 'fdr'). For a full list of tools and packages that are used in this project please refer to the **Supplementary Material**. The user interface is reachable via www.adatlas.org.

3 Results

3.1 Overview of the Alzheimer's disease (AD) Atlas

The AD Atlas is a comprehensive, network-based catalogue of results from large omics studies that is accessible via an interactive, network-based user interface (www.adatlas.org). It was built by inferring relationships between biological entities (e.g. metabolites, SNPs, genes) from large-scale studies, resulting in a highly complex collection of data stored in network format (**Figure 1A**). To increase data accessibility and downstream interpretability, the data was summarized and extracted into a simplified data view (**Section 2.3**). This summarized network representation is accessible through the user interface and consists of four node types; (I) metabolites, mapped across available platforms where possible, (II) genes, including information on associated transcripts, SNPs and proteins (**Supplementary Material**), (III) traits, which describe AD (endo-)phenotypes and biomarkers, and (IV) meta-traits, which are collections of traits loosely following the A-T-(N)-(C) classification [18]. Various different relationship types interconnect these entities (**Supplementary Figure 1B** and **Supplementary Table 1**). Integrating over 20 different studies and analyses, the AD Atlas includes information on 20,353 protein-coding genes, 8,615 proteins, 997 metabolites and 31 unique AD-related traits. Traits include cerebrospinal fluid (CSF) and imaging biomarkers, partially with different covariate settings (adjustment for *APOE* genotype), stratifications or stagings for

neuropathologies, amounting to 59 traits in total (**Supplementary Table 3**). Biological entities and traits are linked by over 1.5 million relationships, representing statistical associations inferred from large-scale quantitative data from population-based cohorts and AD case-control studies. A more detailed summary of the data compiled in the AD Atlas can be seen in **Table 1**.

Table 1. Data compiled in the AD Atlas (simplified data view).

		$N_{entities}$	References
Nodes			
	AD-related meta-phenotypes	10	-
	AD-related phenotypes	59 (31 unique)	-
	Genes (protein-coding)	20,353	
	with differential gene expression data	14,731	[19]
	with differential protein abundance data	3,249	[20]
	Metabolites	997	-
Relationships			
	Genetic associations with AD phenotypes - *traitQTLs*	10,742	this study, [21–25]
	Genetic associations with metabolic traits - *mQTLs*	121,528	this study, [26–30]
	Metabolic associations with AD phenotypes - *mWAS*	66	[31–33]
	Gene co-expression data	465,032	[19]
	Number of brain regions	7	-
	Genetic coregulation - *eQTL*	967,534	[14,34]
	Number of tissues	49	-
	Protein co-abundance data - *partial correlation-based*	13,171	[20,35]
	Metabolic pathways - *partial correlation-based*	1,252	this study, [36]

To enable access to the AD Atlas, we implemented a network- and web-based user interface (**Figure 2**) which allows users to dynamically generate, explore and analyze, context-specific molecular subnetworks surrounding entities of interest i.e. genes, metabolites or AD-related (endo-)phenotypes (**Figure 1B**). These networks provide a rich multi-omics context, integrating multiple layers of information, including eQTLs, pQTLs, mQTLs, gene co-expression and protein co-

abundance, metabolic network reconstructions (based on partial correlations), as well as genetic and differential abundance associations with AD and associated biomarker profiles (schematically seen in **Figure 2D**). They are built using user specified parameters, allowing users to tailor the network to their research question at hand i.e., focusing on a specific brain region or pathway.

Figure 1B showcases the steps of a typical analysis and highlights the core-functionalities of the AD Atlas website. Users first construct a network of interest in a trait-, gene- or metabolite-centric manner via the network settings panel in the network browser (**Figure 2B + C**). The different entry points are defined as follows:

i. <u>Trait- or meta-trait-centric subnetworks</u>

   User specifies trait or collection of traits as input. Genes and metabolites directly associated with the traits of interest are extracted. Furthermore, these entities are annotated with associated genes (genetic association), metabolites (genetic association) and traits (genetic / metabolic association). All relationships between these entities are included in the resulting network. A meta-trait is a collection of traits (see **Supplementary Table 3**).

ii. <u>Gene-centric subnetworks</u>

   User provides gene or set of genes as input. Metabolites and traits that are directly associated with the genes of interest (provided as gene symbol or Ensembl ID) are extracted. Furthermore, all relationships between these entities are included in the resulting network.

iii. <u>Metabolite-centric subnetworks</u>

   User provides metabolite or set of metabolites as input. Traits and genes that are associated with the metabolite of interest (provided as biochemical name) are extracted. Furthermore, all relationships between these entities are included in the resulting network.

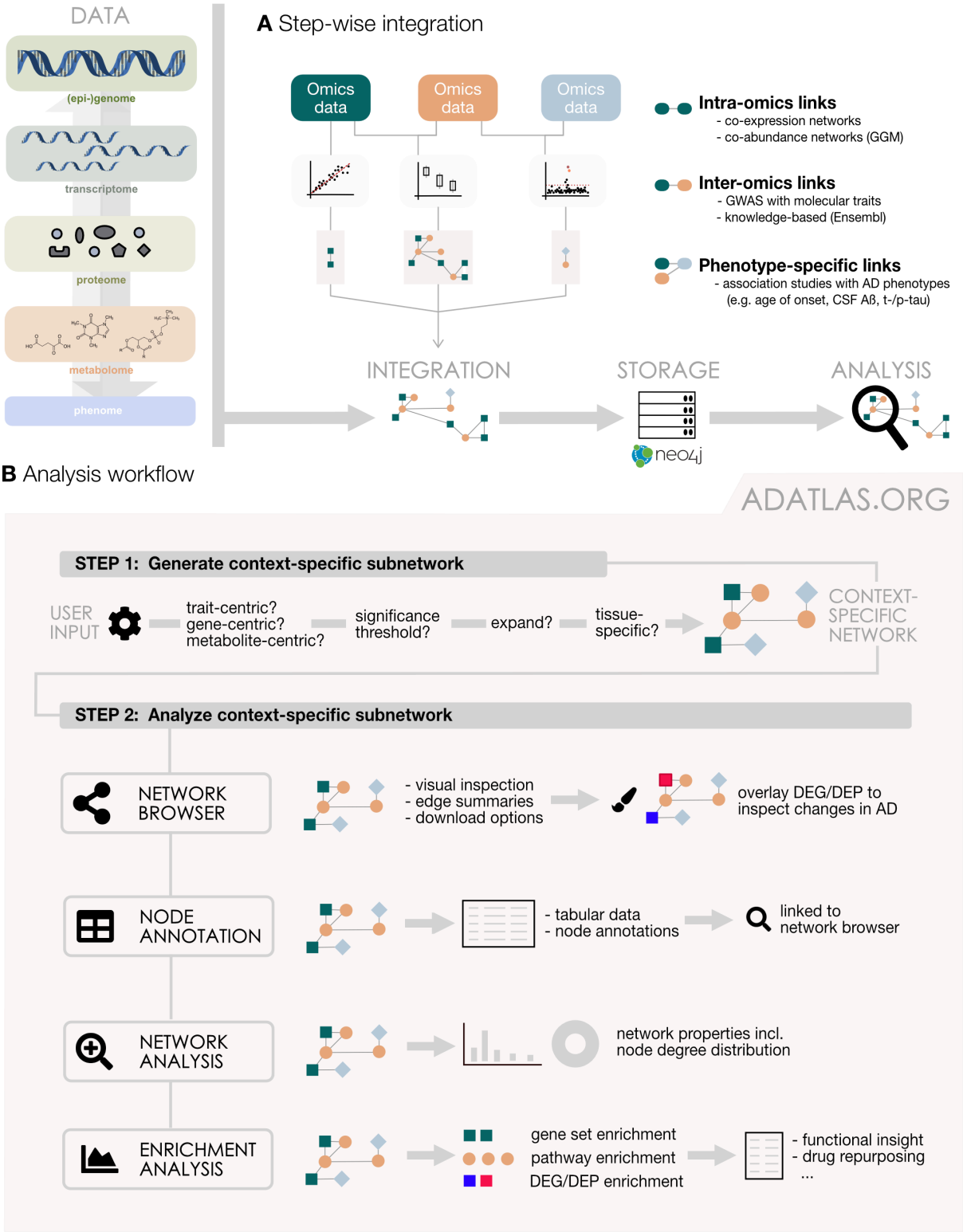


Figure 1. Overview. The AD Atlas is a multi-omics resource that enables the integration and analysis of heterogeneous omics datasets in the context of Alzheimer's disease. **A.** Step-wise multi-omics integration approach underlying the AD Atlas. Using statistical analysis, such as association analysis and partial correlation, omics data collected in large population-based studies is used to infer biological relationships between (inter-omics) and within (intra-omics) omics layers. Links to AD (endo-)phenotypes from large-scale case/control or biomarker studies enable multi-omics exploration in the context of AD. **B.** Users can generate context-specific molecular subnetworks surrounding entities (traits, genes, metabolites) of interest and analyze these networks using tools, such as enrichment analysis, via our interactive user interface (www.adatlas.org; **Figure 2**). DEG: differentially expressed gene; DEP: differentially expressed protein.

Additionally, these molecular subnetworks can be contextualized by applying edge filters (e.g. by tissue or brain region) or adjusting the significance threshold for genetic associations (**Section 2.3**). To gain further insights into the functional neighborhood of metabolites or genes, users can expand input entities before annotation to include the 1-step or 2-step neighbors using co-regulation (eQTL), transcript co-expression or protein co-abundance data for genes, or partial correlation data for metabolites. Once built, the networks can be visually inspected using the network browser (**Figure 2C**) and can be subjected to a number of downstream analyses within the AD Atlas. Information on differentially expressed genes (DEG) and proteins (DEP) can be overlaid onto the networks to investigate the extent and direction of dysregulation in AD using the visualization options (**Figure 2B**). Here, users can choose the underlying association model (sex-specific or pooled analysis) and brain region. Furthermore, the entities in the generated network can be functionally characterized using gene set and pathway enrichment analysis.

In the following two sections, we will showcase how the AD Atlas can be utilized to generate multi-omics subnetworks and provide AD relevant insights for different types of research questions. We focus on two distinct applications; hypothesis-driven drug repositioning, i.e. using established AD pathways to find promising drug candidates, and exploratory analysis, i.e. generating testable hypothesis regarding underlying molecular mechanisms in AD. For this, we built context-specific networks using either genes, clinical diagnosis, or metabolites as input and then assessed the resulting networks with regard to network structure, involvement in AD pathology (association with AD-related phenotypes or evidence of dysregulation at the metabolite, gene or protein level) and gene set enrichment analysis. The subnetworks were constructed by filtering co-regulation edges for the tissue “Brain cortex” and applying a genome-wide significance cutoff, if not stated otherwise. To overlay differential analysis data, we used the setting “AD Diagnosis (males and females)” and tissue “TCX” (temporal cortex) for DEGs and “Control vs. AD” for DEPs. All showcases can be interactively explored under www.adatlas.org/?showcases.

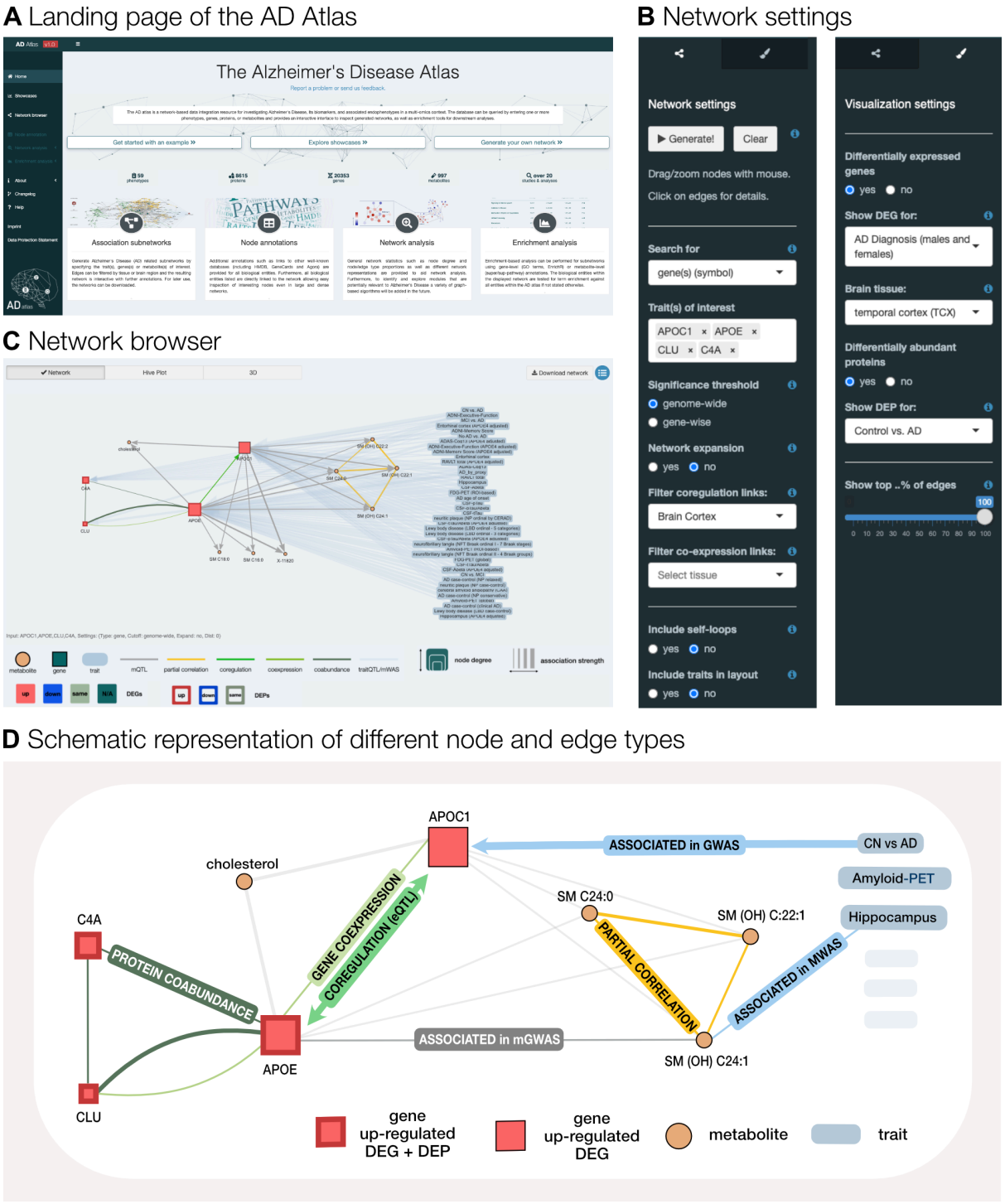


Figure 2. User interface. Users can generate context-specific molecular subnetworks and apply analytical tools via our interactive user interface (www.adatlas.org). **A.** AD Atlas landing page. **B.** In the network browser, users can use the right panel to specify parameters to generate context-specific molecular subnetworks (left). By clicking the paintbrush icon at the top of the panel, additional information, such as differential gene expression in disease, can be overlayed using the visualization options (right). **C.** Network browser showcasing the multi-omics subnetwork surrounding the input genes *APOE, APOC1, C4A* and *CLU*. **D.** Schematic representation of the molecular subnetwork seen in C, showcasing the different node and edge types. For simplicity, not all nodes and edges are included. GWAS: genome-wide association study; mGWAS: GWAS with metabolic traits; MWAS: metabolome-wide association study; protein abundance: partial correlation between proteins; eQTL: expression quantitative trait locus; DEG: differentially expressed gene; DEP: differentially expressed protein.

3.2 Hypothesis-driven applications for drug repositioning approaches

Little progress has been made in regards to effective and disease modifying therapies for Alzheimer's disease [37]. Therefore, to accelerate the path to effective intervention strategies, drug repositioning – the application of available and approved compounds in a novel disease context – has gained increasing attention as a promising alternative to *de novo* drug development [38]. In the following we will highlight the potential of the AD Atlas to advance computational repositioning efforts in AD by integrative analysis of comprehensive multi-omics data. All showcases can be interactively explored at adatlas.org/?showcases.

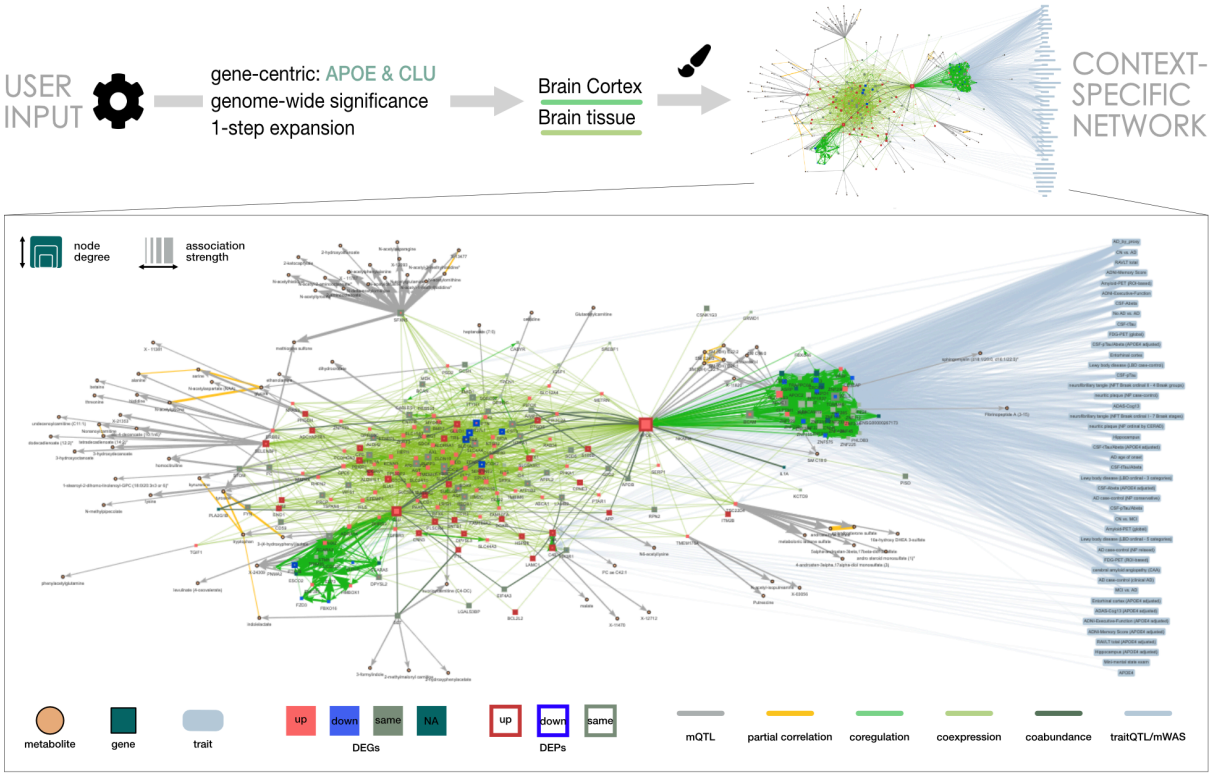
3.2.1 Molecular subnetwork of lipid metabolism and transport identifies known repositioning candidates

In recent years, multiple genetic risk factors for late-onset Alzheimer's disease (LOAD) have been identified through GWAS [10]. The earliest was the discovery of the genetic risk exerted by the ε4 allele of apolipoprotein E (*APOE*) [39], which was followed by the identification of several risk variants in clusterin (*CLU,* also referred to as *APOJ*) [40,41]. Both proteins are involved in lipid metabolism and transport and we hypothesized that these mechanisms could potentially be targeted by available repositioning candidates. We therefore queried the AD Atlas using the two genes as input and expanded to the 1-step functional neighborhood defined by gene co-expression, co-regulation data and protein co-abundance data (**Supplementary Material**). The resulting network provides the molecular context around these AD-associated genes by integrating multiple layers of multi-omics information (**Figure 3A**). As expected, *APOE* displays multiple direct associations with a large number of AD-related phenotypes ranging from disease status (control vs. AD) to CSF and imaging biomarkers. *CLU* is strongly associated with disease status, AD-by-proxy [24] and age of disease onset. The subnetwork contains 218 protein-coding genes in total, which are associated with 81 metabolites. 48 genes harbor SNPs showing a significant association to AD (endo-)phenotypes. Next, we overlaid differential gene expression and protein abundance data from

large-scale case-control studies of results from AMP-AD [19,20]. This was done using the visualization options of the AD Atlas network browser (accessible in the right panel by clicking the paintbrush icon; **Figure 2B**) to further characterize the direction of dysregulation in disease (**Figure 3A**). This revealed an up-regulation on the transcript levels of both *APOE* and *CLU* as well as their functional neighborhood in individuals with AD compared to healthy individuals (133/218 up-regulated (red), 14/218 down-regulated (blue), 60/218 not differentially expressed, 11/218 no data). Protein abundance data showed very similar patterns (33/218 more abundant in AD patients (red border), 6/218 less abundant (blue border), 44/218 no difference in abundance, 135/218 no data).

In order to utilize the potential of such networks towards identification of drug repositioning candidates, we subsequently performed gene set enrichment analysis on all the genes in this subnetwork using molecular drug signatures from the EnrichR database [42] (left hand panel under "Enrichment analysis – Gene set enrichment – enrichR"). Although this is a rather simplistic approach that ignores effect direction or strength, we were able to identify multiple candidate drugs that have been previously proposed and tested in clinical trials among the top hits, i.e. their associated lists of genes affected in drug screens overlapped most significantly with the *APOE/CLU* context network. The obtained list of drugs includes Valproate, a drug with antiepileptic properties, the anti-diabetic drug Rosiglitazone, and Fluoxetine, a selective serotonin reuptake inhibitor (SSRI) (upper table in **Figure 3B**). Unfortunately, clinical trials have failed to show a significant effect on cognition or function for both Rosiglitazone [43,44] and Valproate [45], with the latter also displaying severe toxic effects. SSRIs, especially fluoxetine, showed promising effects on AD pathology in animal models and improved cognition in meta-analysis of short-term human trials of dementia patients with depression [46,47]. However, further large-scale trials are needed to verify if fluoxetine offers benefits in AD patients without depression.

A - Molecular subnetwork surrounding *APOE* and *CLU*



B - Selected gene set enrichment results

Drug Perturbations from GEO 2014 (EnrichR)	*rank*	*overlap*	*q-value**	*OR***	*genes*
Fluoxetine mus musculus gpl1261 gds2803 chdir up	1	17/281	2.62e-06	6.25	APP, TTYH1, FAM107A, SLC1A2, STMN4, AQP4, PTN, ATP1B2, CLU, CNN3, TMEM47, CST3, GJA1, EDNRB, CPE, VIM, ITM2B
Rosiglitazone rattus norvegicus gpl341 adipose tissue gds3850 chdir down	2	17/318	1.20e-05	5.47	APP, TRIL, TMEM176A, HTRA1, LAMC1, NDRG2, S100B, CLU, DHRS3, C4B, SELENBP1, SERP1, EFEMP1, AXL, DPYSL2, ITM2B, DDR2
Rosiglitazone mus musculus gpl1261 control gds4036 chdir down	4	17/342	2.25e-05	5.06	APP, IL33, SCARA5, SDC2, TMEM176A, EPHX1, HTRA1, LAMC1, C4B, CST3, SERP1, AXL, SOX9, VIM, PTGDS, BCL2L2, DDR2
Valproic acid homo sapiens gpl6883 gse26940 chdir up	6	66/3508	1.91e-04	2.06	APP, SRPX, SLC44A3, TRIL, ZNF45, VLDLR, LAMC1, PIK3C2A, GLI3, TMEM47, GJA1, SERP1, HEPH, MDK, DPYSL2, DPYSL3, ZNF404, CPNE3, SOX9, APPL2, WLS, TGIF1, SREBF1, CSNK1G3, SCARA3, SLC2A10, METRN, PHKA1, SORCS2, ACOX2, PLSCR4, DDR2, FTL, SET, CCDC25, HTRA1, ITPR2, SLC1A3, AK3, PBXIP1, PRCP, PIK3R1, PTN, NDRG2, CST3, SELENBP1, TSPAN6, LRIG1, FYN, APOE, KCNN4, PLTP, ABCA1, SCD5, EPHX1, CYBRD1, PRSS35, PC, BCL3, CPE, VIM, SFXN5, PLCD1, BCL2L2, SNTB1, ITM2C
Drug Perturbations from GEO down (EnrichR)	*rank*	*overlap*	*q-value**	*OR***	*genes*
Candesartan DB00796 rat GSE2739 sample 2687	4	21/369	1.63e-07	5.95	APP, SREBF1, SERPINE2, SDC4, SLC1A3, STMN4, AQP4, ATP1B2, NDRG2, S100B, CLU, AGT, CST3, BCAN, GJA1, PHGDH, TMBIM6, APOE, PTGDS, ITM2B, NTSR2
Levetiracetam 5284583 rat GSE2880 sample 2669	7	19/399	1.08e-05	4.87	APP, ENO1, ATP1B2, CLU, AGT, CNN3, GCSH, GLUD1, EDNRB, TRIM35, PHGDH, ACSBG1, TMBIM6, APOE, VIM, PTGDS, ITM2B, NTSR2, ITM2C
Levetiracetam DB01202 rat GSE2880 sample 2777	8	13/181	1.09e-05	7.40	WFS1, STMN4, ENO1, PTN, NDRG2, S100B, CST3, DPYSL2, TRIM35, CPE, APOE, ITM2B, ITM2C

*adjusted p-value using Benjamini-Hochberg **odds-ratio

Figure 3. *APOE/CLU* subnetwork identifies repositioning candidates. A. Multi-omics subnetwork surrounding *APOE* and *CLU* as contained in the AD Atlas. **B.** Gene set enrichment analysis for drug-associated gene expression changes using EnrichR reveals previously proposed candidates Valproate, Fluoxetine and Rosiglitazone among the drugs most significantly affecting the subnetwork (upper table). When focusing on signatures which are opposed to the overall change in AD, we identify Levetiracetam and Candesartan as most promising (lower table).

With the exception of Rosiglitazone, these drugs up-regulate genes in the network. However, the subnetwork surrounding *APOE* and *CLU* displays an up-regulation signature in disease and we hypothesize that drugs that perturb these genes in an opposing manner may be the most promising candidates to exert AD-relevant therapeutic effects [38]. This may also provide a partial explanation for severe side-effects; if the drug and disease perturb gene expression in the same direction this could lead to aggravated symptoms or accelerated disease progression, given that the observed transcriptional changes are not a compensatory mechanism. To test this hypothesis, we repeated the enrichment analysis using a library comprised of only down-regulation signatures (EnrichR, *Drug Perturbations from GEO down*). The results of this analysis can be seen in the lower table of **Figure 3B.**

Interestingly, the top hits include Levetiracetam, a medication that is used to treat epilepsy and is currently being studied to determine whether or not it is able to improve synaptic function and reduce amyloid-induced neuronal hyperactivity as a disease-modifying therapy [48]. With Levetiracetam being investigated in multiple phase II trials (NCT02002819, NCT03489044, NCT03875638) and a low-dose formulation (AGB101) currently being tested in phase II (NCT03461861) and phase III trials (NCT03486938), Levetiracetam is one of the most represented agents among ongoing clinical trials (as of February, 2020) [49]. We further investigated the genes in this subnetwork surrounding *APOE* and *CLU* that are affected by Levetiracetam. To this end, we queried the AD Atlas using the genes listed in **Figure 3B** (via URL query parameters; **Supplementary Figure 4A**). This analysis revealed that of the 20 genes, 15 show a dysregulation at the mRNA level with 14 in opposing direction to the drug (up-regulated in AD) and only one showing down-regulation.

Another interesting candidate that was identified by this analysis is Candesartan, an angiotensin receptor blocker typically used for the treatment of hypertension. Of the 21 genes that are affected

by the drug in the subnetwork, more than half also display a perturbed transcriptional signature in AD (13 up-regulated and one down-regulated in the TCX of AD patients; **Supplementary Figure 4B**). Candesartan has recently been studied in a phase II trial (NCT02646982) to investigate its effect on individuals with mild cognitive impairment that are positive for AD biomarkers. Although, no results have been published for this trial, previous clinical trials suggest beneficial neurocognitive effects following Candesartan treatment in older individuals with hypertension and mild cognitive impairment [50,51]. Of note, investigation of the network structure in the AD Atlas revealed that genes down-regulated by Levetiracetam and Candesartan lie within distinct genetic loci and are connected at the mRNA and protein level (**Supplementary Figure 4**). Furthermore, both affect the AD-risk genes *APOE* and *CLU*, which were provided as input, as well as *APP.*

In summary, we were able to show that the AD Atlas is able to identify plausible repositioning candidates using a simple enrichment approach and requiring minimal analysis steps. The top hits were enriched with candidates that are either being tested or have been tested in clinical trials. After additional analysis using the AD Atlas, we propose Levetiracetam and Candesartan as most promising candidates as their associated list of genes affected in drug screens overlaps significantly with the molecular context network surrounding *APOE* and *CLU* and they affect disease-perturbed genes in an opposing manner.

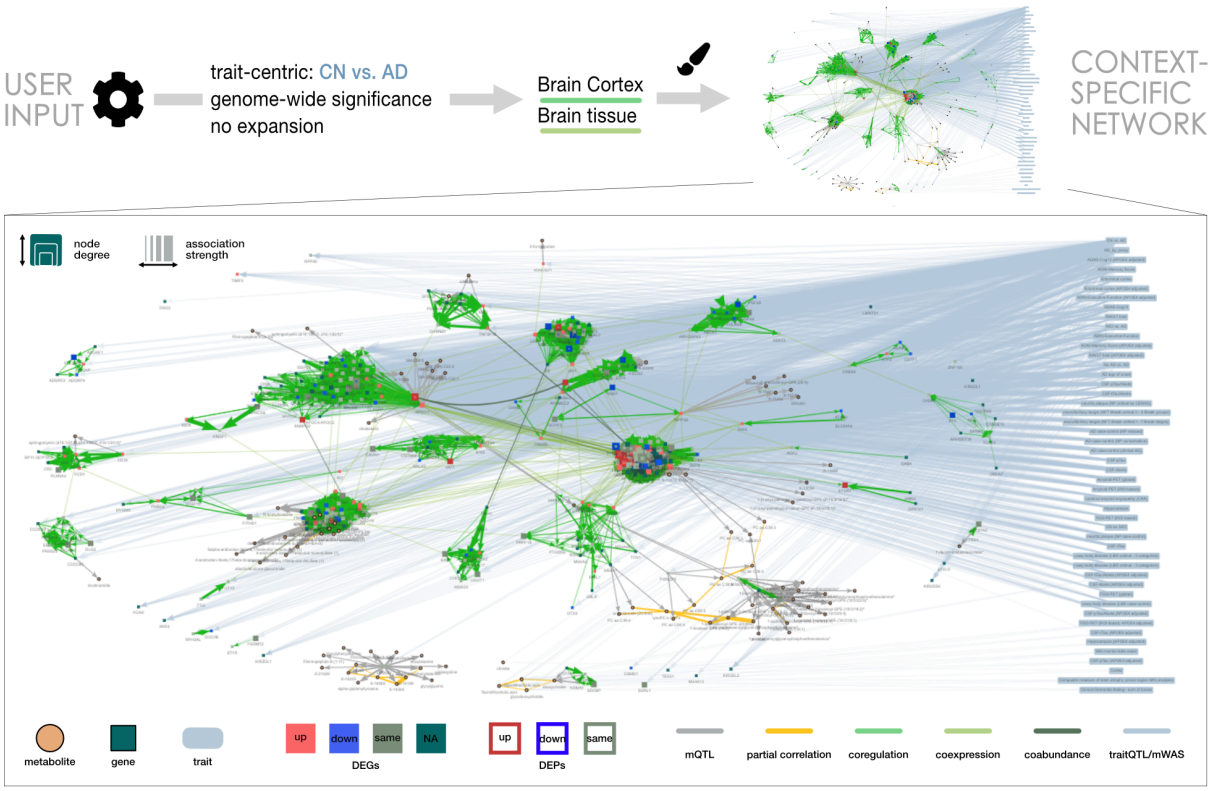
3.2.2 Disease-associated molecular subnetwork provides global view on AD

The previous analysis uses a user-defined set of genes as starting-point for the repurposing analysis, allowing an in-depth and focused view on specific aspects (lipid metabolism) of the disease. To move beyond that, the AD Atlas also provides a trait-specific entry-point which enables the generation of global, data-driven views on AD and its associated (endo-)phenotypes, enabling more comprehensive analyses. Here, entities (genes and metabolites) associated with a specific (set of) AD-related trait(s) identified in large-scale GWAS and MWAS, are embedded into their multi-omics

context that is annotated with metabolic and genetic associations as well as intra-omics links. Using this entry-point, we repeated the analysis steps outlined in the previous example. We selected the results of large-scale AD case-control studies from IGAP [23,25] and ADNI (trait “CN vs. AD”) as input to build a global molecular subnetwork of genetic risk for AD. The network can be seen in **Figure 4A** and consists of 380 genes and 91 metabolites. The network shows multiple tightly connected clusters of genes, which are indicative of the underlying genetic architecture (genetic loci are tightly connected through co-regulation edges). To investigate the degree and direction of perturbation at the mRNA and protein level, we used the AD Atlas visualization options to overlay results of differential gene and protein analysis. This revealed both up-regulation (70 transcript-level, 10 protein-level) and down-regulation (52 transcript-level, 10 protein-level) of the AD-related subnetwork.

Utilizing the previously described concept of gene set enrichment to identify repositioning candidates, we perform the analysis using the drug perturbations library from GEO via EnrichR. The only two drugs that display a significant overlap of genes with the trait-centric network after multiple testing correction are Letrozole and Rosiglitazone (upper table in **Figure 4B**). Unfortunately, both of these agents do not seem to be promising candidates. Letrozole is an aromatase inhibitor used to treat estrogen receptor–positive breast cancer in post-menopausal women and aromatase inhibitors have been associated with adverse effects including negative effects on cognition and potential long-term neural effects [52]. Furthermore, Rosiglitazone, as discussed previously, has failed to show significant effects on cognition in clinical trials [43,44]. Therefore, we focused the enrichment analysis again on drug induced signatures of down-regulated gene expression (gene set *Drug Perturbations from GEO down* from EnrichR).

A - Multi-omics contextualization of large-scale AD case-control GWAS and MWAS



B - Selected gene set enrichment results

Drug Perturbations from GEO 2014 (EnrichR)	*rank*	*overlap*	*q-value**	*OR***	*genes*
Letrozole homo sapiens gpl3921 gse33366 chdir down	1	22/404	5.74e-03	3.10	DDR1, EPHX2, HLA-B, B3GALT4, HLA-C, TAP1, IFI35, CLU, HLA-E, PSMB9, C4B, C4A, AZGP1, BCAM, BCL3, FLOT1, TIMP3, LY6G5C, CFB, HLA-DRB1, CD55, ZKSCAN1
Rosiglitazone homo sapiens gpl570 gds2453 chdir down	2	21/403	8.72e-03	2.95	VAT1, CR1, GPX4, MS4A7, HLA-B, HLA-C, NR1H3, TREM2, DTX4, AIF1, MS4A4A, MS4A6A, SDCBP, HLA-DRA, APOE, CD46, CD68, CLIC1, HLA-DRB1, HSPA1B, PICALM
Drug Perturbations from GEO down (EnrichR)	*rank*	*overlap*	*q-value**	*OR***	*genes*
Etanercept DB00005 human GSE36177 sample 2596	5	14/223	1.42e-02	3.56	HLA-DRB5, MS4A7, KIR3DL1, AOAH, KIR2DL1, PRRC2A, MS4A6A, HLA-DMA, HLA-DMB, HLA-DRA, UGGT1, HLA-DRB1, CD55, HSPA1B
Citalopram DB00215 human GSE7036 sample 3292	7	16/288	1.67e-02	3.14	HLA-DRB5, PVRIG, HLA-B, HLA-C, IFI35, PSMB8, HLA-E, HLA-DMA, HLA-DMB, STAG3, HLA-DRA, HLA-DQA1, HLA-DRB1, HSPA1B, HLA-DPA1, HLA-DQB1

*adjusted p-value using Benjamini-Hochberg **odds-ratio

Figure 4. Alzheimer's disease case-control subnetwork identifies potential repositioning candidates. A. Multi-omics contextualization of large-scale AD case-control GWAS and MWAS studies. The network was built using the trait "CN vs. AD" as input for trait-centric subnetwork generation. **B.** Gene set enrichment analysis for drug-associated gene expression changes in drug perturbation gene sets from GEO via EnrichR reveals Letrozole and Rosiglitazone as the only drugs significantly affecting the subnetwork. By focusing on down-regulation signatures, we obtain more significant hits, among which Etanercept and Citalopram seem to be the most promising candidates.

This analysis yielded multiple significant hits (lower table in **Figure 4B**), including Citalopram, a selective serotonin reuptake inhibitor used in the treatment of depression. Using the AD Atlas, we further investigated whether the genes perturbed by this agent are also perturbed in AD and if its effects on transcript levels are in opposing direction. Citalopram affects a total of 16 genes in the trait-centric subnetwork and more than half are differentially expressed in AD (nine up- and one down-regulated). Furthermore, Escitalopram, the (S)-stereoisomer of Citalopram, is currently being studied as a treatment for agitation in AD patients in a phase III trial (NCT03108846) and has also entered a phase I trial as a cognitive enhancer (NCT03274817).

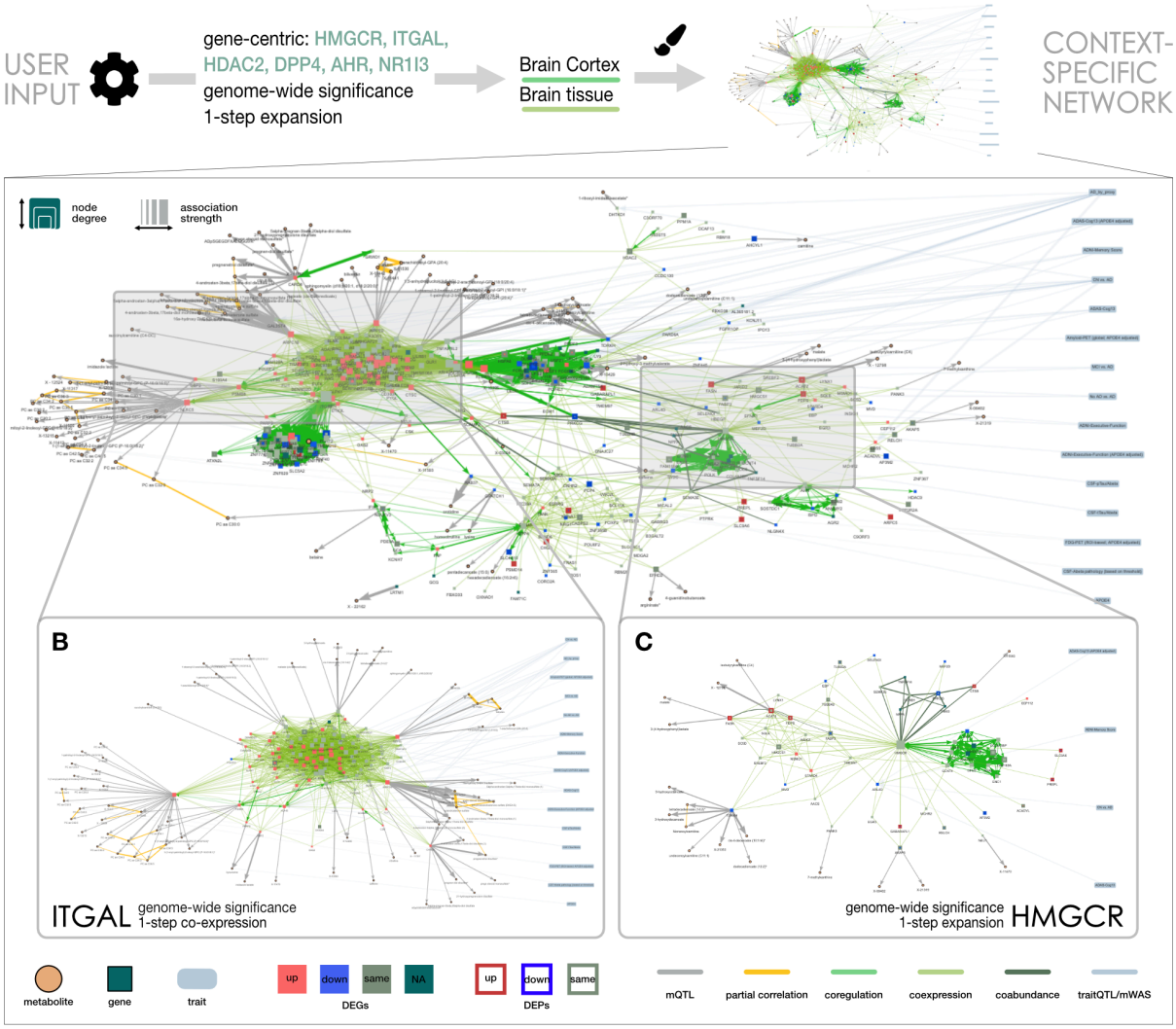
Another interesting finding was Etanercept, a tumor necrosis factor α (TNFα) inhibitor that is used to treat autoimmune diseases (lower table in **Figure 4B**). Elevated levels of TNFα, an inflammatory cytokine, in the brain have been linked to AD and proposed as a potential therapeutic target [53]. Repeating the analysis steps outlined for Citalopram, we find that the Etanercept gene set has 14 overlapping genes of which five show up-regulation in AD. It is interesting to note that the subnetwork also shows a significant overlap with genes that are up-regulated by Etarnecept, indicating that the drug may target both compensatory and disease mechanisms. Furthermore, there is some minor overlap between up- and down-perturbation gene sets which may point to a degree of variability within the response. Safety and tolerability of Etarnecept in AD has been established in a small randomized, placebo-controlled, double-blind phase II trial (NCT01068353) [54] but the drug has not been studied in a phase III trial, despite multiple large observational studies indicating a reduced risk of AD among patients treated with Etarnecept [55]. Further supporting evidence is provided by a second generation TNFα inhibitor, selective for the soluble form of TNFα, that has also shown promising results in preclinical studies [56,57] and was recently investigated in a Phase I trial (NCT03943264). Both Etarnecept and Citalopram seem to affect the immunoregulatory human leukocyte antigen (HLA) complex, which has been implicated in neurodegenerative diseases, including AD [23,58].

In conclusion, the AD Atlas enables the generation of global, disease-related molecular networks without the need to perform complex data analysis. These networks can be used in downstream analysis within the web-based user interface, for instance to identify plausible repositioning candidates using gene set enrichment. In addition, the networks and lists of contained entities can be directly downloaded for offline follow-up analyses. Despite coming from diverse sources and cohorts, the totality of data collected in the AD Atlas can provide valuable insights into the underlying mechanisms of disease, as we demonstrate with the identification of Etarnecept and Citalopram as promising repositioning candidates through potential modulation of neuroinflammatory pathways.

3.2.3 Statin target ITGAL links to neuroinflammation through TREM2 signaling

Statins are a class of lipid-lowering drugs that are used to reduce the risk of cardiovascular diseases, such as atherosclerosis and peripheral artery disease. Statins exert their primary therapeutic effect by inhibiting the rate-limiting enzyme, 3-hydroxy-3-methylglutaryl-coenzyme A reductase (HMGCR), in the cholesterol producing mevalonate pathway. In addition statins have been associated with a wide range of secondary effects [59,60]. Observational studies have reported a possible association between statin use and reduced risk of Alzheimer's disease [61–63], although the evidence has been inconsistent with reported differences between patient subpopulations as well as individual statin drugs [64–66]. To investigate this potential link to AD pathophysiology we used the drug targets of statins, as annotated in DrugBank [67] (**Figure 5E**) and constructed and annotated their molecular context network in the AD Atlas (**Figure 5A**). To this end, we used *HMGCR, ITGAL, HDAC2, DPP4, AHR* and *NR1I3* genes as input, added their 1-step co-expression, co-regulation and protein co-abundance neighbors by network expansion and colored genes according to their differential expression in AD patients using the visualization options of the AD Atlas. Of the 310 genes in the resulting network, 24 genes harbor SNPs showing a significant association to AD (endo-)phenotypes and over one third (n=120) show differential expression at the transcript level in the temporal cortex of AD patients.

A - Molecular subnetwork surrounding statin targets annotated in DrugBank



D - Top 5 gene set enrichment results

WikiPathway 2021 human (EnrichR)	*overlap*	***q-value**** *(ITGAL**)*	*OR****	*genes*
TYROBP causal network in microglia WP3945	17/61	**1.04e-14** (2.08e-23)	25.91	RBM47, IGSF6, PLEK, ITGB2, TCIRG1, SLC1A5, TGFBR1, PYCARD, APBB1IP, CD4, SLC7A7, ADAP2, BIN2, ELF4, CD37, GAL3ST4, TMEM106A
Cholesterol Biosynthesis Pathway WP197	7/15	**1.86e-07** (-)	56.84	FDPS, SQLE, HMGCS1, SC5D, MSMO1, MVD, HMGCR
Cholesterol metabolism (includes both Bloch and Kandutsch-Russell pathways) WP4718	10/46	**1.86e-07** (-)	18.20	FDPS, SQLE, EBP, HMGCS1, FASN, SC5D, MSMO1, MVD, HMGCR, SREBF2
Microglia Pathogen Phagocytosis Pathway WP3937	9/40	**6.66e-07** (1.86e-09)	18.96	HCK, FCER1G, ARPC1B, ITGB2, RAC2, TREM2, PTPN6, PIK3CG, VAV1
Mevalonate pathway WP3963	5/7	**9.82e-07** (-)	161.38	FDPS, HMGCS1, MVD, HMGCR, ACAT2

*adjusted p-value using Benjamini-Hochberg **enrichment for 1-step co-expression network surrounding ITGAL ***odds-ratio

E - Statins and their targets

	HMGCR	ITGAL	HDAC2	DPP4	AHR	NR1I3
Rosuvastatin	■	■				
Lovastatin	■	■	■			
Simvastatin	■	■	■			
Fluvastatin	■		■			
Atorvastatin	■		■	■	■	■
Pitavastatin	■	■				
Pravavastatin	■		■			

Figure 5. Subnetwork of statin targets link to TYROBP signaling. A. Multi-omics subnetwork surrounding the statin targets annotated in DrugBank; *HMGCR*, *ITGAL*, *HDAC2*, *DPP4*, *AHR* and *NR1I3*, as well as a co-expression and co-abundance subnetwork only surrounding (**B**) *ITGAL and* (**C**) *HMGCR*. **D.** Gene set enrichment analysis using EnrichR (*WikiPathway* gene set). **E.** Statins taken from DrugBank and their annotated targets.

Visual inspection of the network indicates two tightly connected clusters surrounding *ITGAL*, one driven by co-regulated genes (dark green edges) and one driven by co-expressed genes (light green edges). The latter is predominantly up-regulated in disease, with an exception of *POU2F2* which is down-regulated. Another co-regulation cluster is formed by co-regulated genes surrounding *NR1I3*, some genes of which (*ARHGAP30, FCGR2A, FCGR3A, FCER1G*) are tightly linked with the *ITGAL* co-expression module. Both co-regulation clusters contain multiple genes that are dysregulated in AD. All target genes either show a dysregulation at the transcript or protein level in AD through direct evidence or via their associated neighborhood. For example, the primary statin target *HMGCR* (**Figure 5C**) does not show any evidence of differential regulation or association to AD phenotypes. However, of its 41 co-expressed genes and co-abundant proteins in the temporal cortex, 19 protein-coding genes show either up- or down-regulation at the transcript level and nine at the protein level, suggesting a functional involvement of this pathway in AD.

Next, we used the gene set enrichment analysis within the AD Atlas to functionally characterize the pathways targeted by statins. Here, we found a significant enrichment for the TYROBP causal network (EnrichR, WikiPathways, P_{FDR} = 7.89e-15), an immune- and microglia-specific module that has been implicated in late-onset Alzheimer's disease (LOAD) [68]. This enrichment is driven by the co-expression network surrounding *ITGAL* (**Figure 5B**), also known as *CD11a*, a subunit of the integrin leukocyte function associated antigen-1 (LFA-1) which is involved in a variety of immune-related functions [69]. Interestingly, only a subset of statins (Rosuvastatin, Lovastatin, Simvastatin, Pitavastatin) target *ITGAL* which may explain the heterogenous results reported in studies. Furthermore, the co-expression neighborhood of *ITGAL* shows substantial dysregulation in disease with 50 of 103 genes differentially expressed in AD (49 up-regulated, one down-regulated) and 11 genes harboring disease associated SNPs. TYROBP (Dap12) is an adaptor molecule involved in the transduction pathway of *TREM2* as well as *CD33*, a known AD risk gene [70,71], and CR3 (*ITGAM* and

ITGB2). Both *TREM2* and *ITGB2* are up-regulated in disease and contained in the network shown in **Figure 5B**.

Our analysis showcases how the AD Atlas can be utilized to investigate candidate drugs in an AD context. The identified functional link to TYROBP signaling indicates that statins targeting *ITGAL* may exert potential protective effects [72,73] through modulation of neuroinflammatory pathways [74].

3.3 Application examples for exploratory analysis

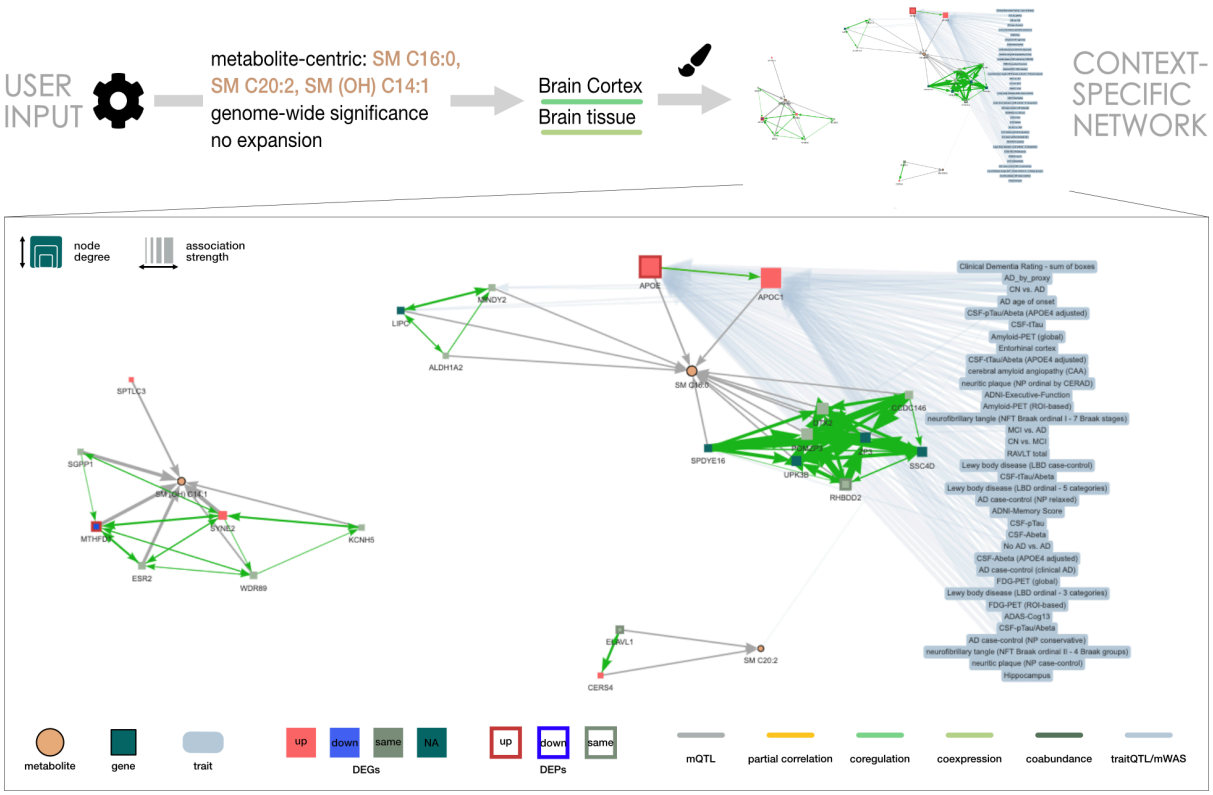
High failure rates of AD drugs in clinical trials have emphasized our incomplete understanding of crucial biological aspects of this complex disease [37]. Basic research investigating the underlying molecular disease mechanisms and mapping the trajectory of biochemical changes in AD, will therefore be crucial to identify novel therapeutic targets and guide future drug development efforts. In the following examples we showcase how the AD Atlas can be used to investigate and formulate new AD-related hypothesis. We focus on the contextualization of metabolic and immune-related AD findings, although there are many more areas of application.

3.3.1 Contextualization of links between the sphingomyelin pathway and AD pathology

In a previous study, we identified sphingomyelin species (SMs) of differing lengths to be implicated in early vs. late stages of AD [75]. More precisely, we found SM C16:0 to be associated with CSF $A\beta_{1-42}$ pathology, while SMs with longer fatty acid chains (≥C20) were correlated with brain atrophy and cognitive decline. This study identified three SMs associated with AD, labeled as SM (OH) C14:1, SM C16:0, SM C20:2 (of note, these vendor-specific labels will be updated to better reflect the currently accepted notations for SMs in future releases). We generated a metabolite-centric subnetwork via the AD Atlas user interface to contextualize these findings and gain a better understanding of their potential functional role in AD.

The generated subnetwork is shown in **Figure 6A.** All SMs are associated with at least two genes at a genome-wide significance level. Interestingly, these particular SMs are not directly linked to each other through partial correlations. Also, while the associated genes are partially interlinked by co-regulation edges pointing to the same genetic loci, the individual "SM-gene-clusters" are not

A - Molecular subnetwork surrounding SM (OH) C14:1, SM C16:0, SM C20:2



B - Top 5 gene set enrichment results

Reactome 2016 (EnrichR)	*overlap*	*q-value**	*OR***	*genes*
Sphingolipid de novo biosynthesis Homo sapiens R-HSA-1660661	3/33	2.53e-04	104.99	CERS4, SPTLC3, SGPP1
Sphingolipid metabolism Homo sapiens R-HSA-428157	3/74	1.46e-03	44.27	CERS4, SPTLC3, SGPP1
Chylomicron-mediated lipid transport Homo sapiens R-HSA-174800	2/17	2.13e-03	133.09	LIPC, APOE
Lipoprotein metabolism Homo sapiens R-HSA-174824	2/34	5.20e-03	62.33	LIPC, APOE
Metabolism of lipids and lipoproteins Homo sapiens R-HSA-556833	5/659	5.20e-03	8.69	CERS4, LIPC, SPTLC3, APOE, SGPP1

*adjusted p-value using Benjamini-Hochberg **odds-ratio

Figure 6. Contextualization of metabolomics-guided insights points to SM de novo biosynthesis. A. Multi-omics subnetwork surrounding three sphingomyelin (SM) species (SM (OH) C14:1, SM C16:0, SM C20:2) that are altered in biomarker defined stages of AD [75]. **B.** Gene set enrichment analysis using the *Reactome* gene set accessible via EnrichR identifies three genes (*CERS4, SPTLC3 and SGPP1)* involved in SM de novo biosynthesis.

interlinked. These aspects may point to their involvement in distinct pathways or pathway steps. Furthermore, 35 of the 36 AD-related traits are associated to SM C16:0 related genes, which include AD risk genes *APOE* and *APOC1*. The SM (OH) C14:1 component shows no direct or indirect trait associations, while SM C20:2 is directly associated with the clinical dementia rating - sum of boxes score. Overlay of differential gene and protein expression data using the visualization options provided by the AD Atlas reveals an up-regulation at the mRNA level of 5 of 22 genes (APOE locus (*APOE*, *APOC1*), *SPTLC3*, *SYNE2* and *CERS4*) and down-regulation of one gene (*MTHFD1*), which displays an opposing effect at the protein level (up-regulation).

To further characterize the potential functional involvement of these SM-associated genes in AD, we performed an enrichment analysis using the *Reactome 2016* library via EnrichR. This identifies three genes (*CERS4, SPTLC3 and SGPP1)* involved in SM *de novo* biosynthesis (P_{FDR} = 2.53e-04, **Figure 6B**). Interestingly, these genes have previously been identified in a study that involved multiple, time-intensive manual mapping steps [76]. Here, the genes were categorized into two functional categories: global sphingomyelin synthesis (*SPTLC3*, *CERS4*) and synthesis and degradation of sphingosine-1-phosphate (*SGPP1*), highlighting a possible role for sphingosine-1-phosphate and its receptors in AD pathogenesis. AD mouse models indicate a potential benefit of Fingolimod, an FDA-approved S1P analog used for the treatment of multiple sclerosis [77–79]. Furthermore, long-term Fingolimod treatment in multiple sclerosis patients has showed positive effects on cognition [80]. Therefore, Baloni et al. applied a drug repositioning approach by treating APP/PS1 mice with Fingolimod, finding that prolonged S1P pathway modulation can rescue both the proposed cellular mechanism of hippocampus-related memory and cognitive deficits in these mice, further supporting this pathway as a high priority target for AD [76].

In conclusion, this analysis highlights the ability of the AD Atlas to contextualize hypotheses or findings from previous studies and to thereby point to novel, AD-related insights without the need

for bioinformatics resources or time-intensive manual analyses. Here, we used three sphingomyelin species that have previously been associated with AD endophenotypes [75], to further investigate their involvement in mechanisms of disease. We find a link to the sphingosine-1-phosphate pathway, replicating results of a recent study [76].

3.3.2 Subnetworks surrounding marker genes for homeostatic microglia and disease-associated microglia suggests possible involvement of blood androgens

Genomic analyses in AD and animal models of the disease identified a specific activation program which drives the transition from homeostatic microglia to disease-associated microglia [81]. However, the molecular mechanisms underlying this transition remain incompletely understood and the extent to which this process involves AD susceptibility genes has not been assessed in an integrated fashion. We used the AD Atlas to identify gene modules defining homeostatic vs. disease-associated microglia and to evaluate their respective links with large-scale genetics, proteomics and metabolomics data. Two canonical gene markers of homeostatic vs disease-associated microglia, namely *TMEM119* [82,83] for homeostatic microglia and *TREM2* [84] for disease-associated microglia, were entered as single query genes. We further analyzed the corresponding co-expression networks identified in brain tissue using the AD Atlas with regard to: i) identity and number of genes harboring SNPs previously associated with AD susceptibility and/or AD-related traits in GWAS studies, ii) identity and number of genes exhibiting increased levels in AD brains at the transcript and/or protein levels, iii) identity and number of genes harboring SNPs previously associated with metabolite concentrations in healthy individuals. An overview of the analysis steps can be seen in **Figure 7A**.

We identified a total of 55 genes being co-expressed with *TMEM119* but not *TREM2* in AD tissue, of which five genes are linked to AD-associated SNPs: *ARPC1B, TMEM106A, INPP5D, HLA-DMA, ARHGAP45* (**Supplementary Table 4**). Out of these five genes, two (*ARPC1B, INPP5D*) harbor mQTL

associations with bilirubin, biliverdin, 1-archidonoyl-GPA (20:4), unknown metabolites and multiple androgenic steroids, including epiandrosterone sulfate, androsterone sulfate and dehydroisoandrosterone sulfate (DHEA-S) (**Supplementary Table 5**). Additionally, both genes show higher gene expression levels in AD (temporal cortex, AD vs. Control). The overlapping genetic associations with AD and levels of metabolites from the androgen pathway in the *ARPC1B* locus is of particular interest. Androgens are a class of sex steroid hormones that are responsible for the development of male sex characteristics [85] and also play important roles in female reproductive function [86]. Females have a higher susceptibility to AD [87] and studies have linked age-related depletion of the androgen testosterone to an increased risk of AD in men [88,89]. Since *ARPC1B* has been linked to the branching and motility of microglial ramifications [90], this might suggest a potential molecular relationship between androgen levels, ageing and the ability of microglia to extend ramifications in the context of AD.

We also identified 64 genes co-expressed with *TREM2* but not *TMEM119* in brain tissue, of which nine genes map to AD-associated loci: *APOC1, ITGAX, LST1, HLA-DPB1, GAL3ST4, ITGAM, SPI1, AIF1, HLA-DRB1* (**Supplementary Table 6**). In comparison to the TMEM119 module, the TREM2 module contains a larger number of genes with overlapping AD associations, including the known genetic risk factor *APOC1* which harbors multiple SNPs showing strong associations to a multitude of AD phenotypes (36 in total). Furthermore, while both TMEM119- and TREM2-specific networks show an overall up-regulation at the transcript level (temporal cortex), the TMEM119 network only contains one gene where this up-regulation translates to the protein level (*CNPY3*). The TREM2 subnetwork in contrast contains four such genes (*GPX1*, *CAPG*, *COTL1*, *NPC2*), possibly indicating a higher dysregulation at the functional level. Taken together, this supports a major role for *TREM2* in the engagement of microglia toward a disease-associated phenotype. From the nine genes genetically associated with AD, five (*A1F1, APOC1, GAL3ST4, HLA-DRB1, LST1*) also harbor mQTL associations with cholesterol, multiple SMs, unknown metabolites and multiple androgenic steroids (**Supplementary Table 7**) and two (*APOC1, HLA-DRB1*) show higher gene expression in the temporal

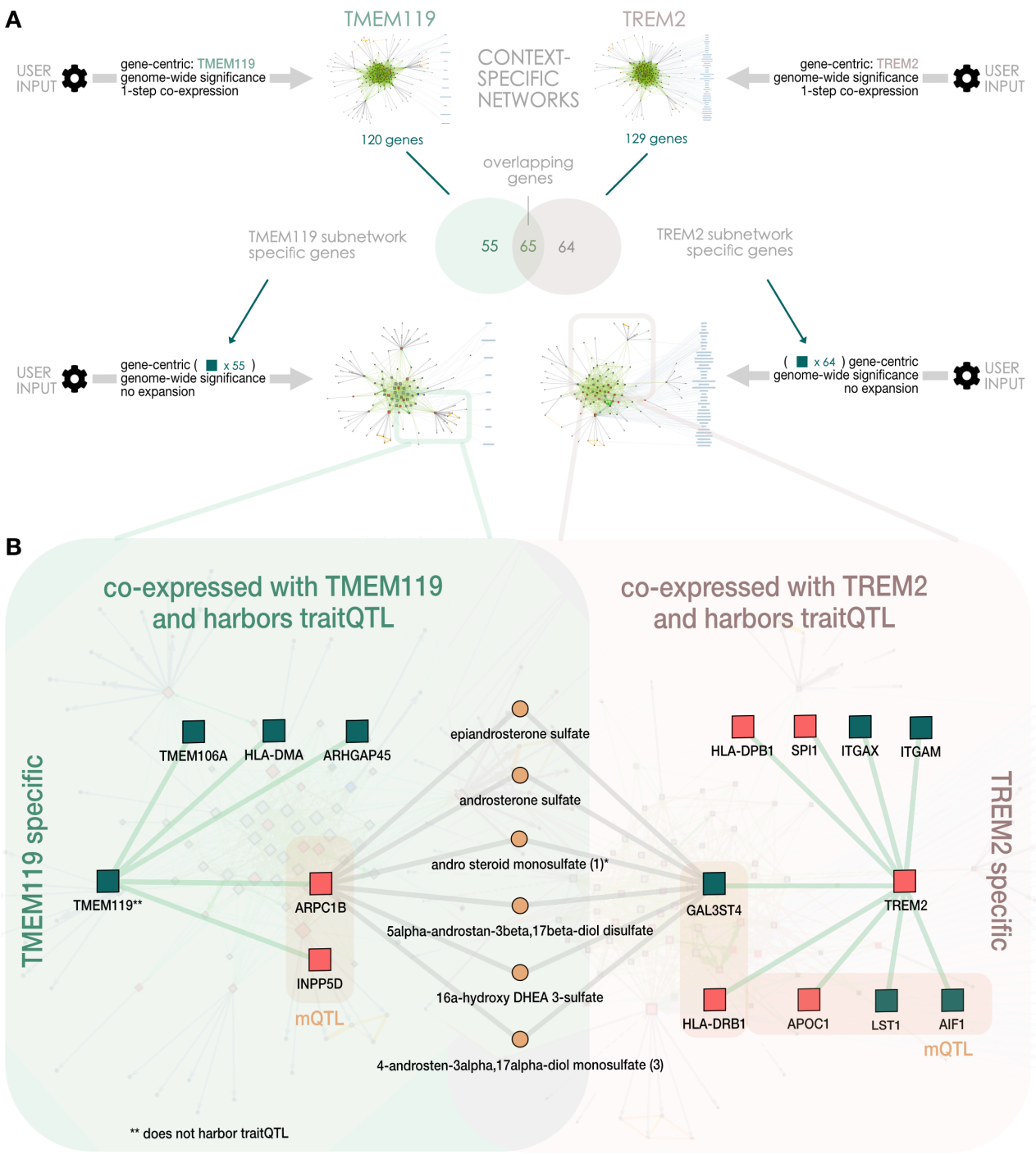


Figure 7. Investigating the transition between homeostatic microglia and disease-associated microglia (DAM). A. Overview of the analysis. First, gene co-expression networks were built around marker genes for homeostatic (*TMEM119*) and DAM (*TREM2*). Then gene modules containing genes co-expressed exclusively with one of these genes were constructed and analyzed. **B.** Schematic representation of the marker-specific subnetworks. Genes that show co-expression in the brain with *TMEM119* but not *TREM2* and harbor AD-related SNPs are shown on the left (green background). Those that also contain mQTLs are highlighted (orange). Genes that are significantly upregulated are highlighted in red. Genes that show co-expression in the brain with *TREM2* but not *TMEM119* and harbor AD-related SNPs are shown on the right (red background). Those that also contain mQTLs are highlighted (orange). Androgen steroids associated with both *ARPC1B* and *GAL3ST4* are depicted to highlight this overlap. It is important to note that only selected relationships are shown. For a detailed version of these networks please refer to our website (www.adatlas.org/?showcases).

cortex in AD. As observed in the *TMEM119* module, one gene of the *TREM2* module, namely *GAL3ST4*, is associated with androgen metabolites. Although few data are available regarding the role of *GAL3ST4* in microglia, the gene was reported to be part of a *TYROBP* brain-expressed gene module crucially involved in the development of LOAD [68]. Interestingly, a metabolite-centric search for the direct subnetwork surrounding the androgen steroids associated with both *ARPC1B* and *GAL3ST4* (overlap of metabolites is seen in **Figure 7B**) revealed an association of these metabolites with another AD-specific risk locus; *ZCWPW1/NYAP1/PILRA* (7q22.1) [23,25,91], which has been linked to myeloid enhancer activity, microglia function and neuroinflammation [91,92].

Lastly, to see if these networks can again point to drug compounds targeting the activation of microglia, we repeated the gene set enrichment analysis described previously using the 1-step co-expression network surrounding both *TMEM119* and *TREM2*. This identified Fasudil (P_{FDR} = 7.02e-16, EnrichR, *Drug Perturbations from GEO down*) among the top significant hits. Fasudil is an inhibitor of Rho-kinase (ROCK) and approved for the treatment of cerebral vasospasm in Japan. Post mortem data suggests that ROCK protein levels are elevated in AD brains [93] and preclinical data from *in vitro* and *in vivo* studies, including animal models of AD, indicate that Fasudil may be able to reduce the burden of tau protein [94] and promote an anti-inflammatory microglial phenotype [95,96]. Currently, two ongoing clinical phase II trials are investigating the use of Fasudil in tauopathies (NCT04734379) and the effectiveness of an oral formulation of Fasudil in patients with dementia (NCT04793659). Interestingly, a previous study reported experimental evidence linking androgen levels through androgen receptor signaling to the levels of miRNA-135a which targets ROCK, providing a potential mechanistic model integrating the different omics entities contained in the AD Atlas-derived subnetwork [97].

Overall, our analyses point to a potential involvement of blood androgens in the transition from homeostatic to disease-associated microglia during the course of AD. Using gene set enrichment we

were able to identify Fasudil, an inhibitor of ROCK, as a promising drug repositioning candidate. We hypothesize that age-related decreases of androgen levels may result in an upregulation of ROCK via androgen-mediated pathways, leading to microglial activation and neuroinflammation [98,99]. This exploratory analysis highlights how the AD Atlas can inform testable hypothesis, such as the potential link between androgen signaling, ROCK activity and microglial activation in AD, that can be investigated in follow-up experiments.

4 Discussion

We here present the AD Atlas, a network-based resource that consolidates experimental data from large cohort studies, building a global molecular view of AD. Integrating data in a largely hypothesis-free and data-driven manner, the AD Atlas enables researchers to explore the disease in a multi-omics context starting from any genes, metabolites or phenotypes of interest. Through a browser-based interface to the AD Atlas network, we provide researchers of varying backgrounds and interests an easy-to-use resource for AD knowledge discovery.

The molecular network underlying the AD Atlas is derived using large-scale population-based data and genomic annotation databases and provides a generalized framework to study multi-omics relationships globally. Using association data from large-scale studies on diverse aspects of AD, including hundreds to hundreds of thousands of individuals, this framework is then transformed into an integrated multi-omics knowledgebase for markers of AD. These data include recent large genetic association (meta-)analyses of AD and AD biomarkers from NIAGADS and other large efforts, as well as differences in transcriptomics, proteomics, and metabolomics markers observed in AD or in relation to AD endophenotypes. The latter are predominantly based on data generated on thousands of human blood and brain tissue samples from multiple brain regions using state-of-the-art technologies and analyzed using standardized processing pipelines through the AMP-AD program and partnering initiatives. By integrating multiple independent datasets of the same analysis type

(e.g. GWAS on the same traits conducted in different cohorts), we build additional data confidence by independent replication and make this data available to the user.

Through its network-based user interface, available at www.adatlas.org, the AD Atlas enables users to conduct flexible and context specific analyses tailored to the research question at hand, without the necessity of local bioinformatics capacities. We provide multiple-entry points, allowing researchers to dynamically generate multi-omics subnetworks surrounding a gene, metabolite or trait of interest. Thereby, the analysis is not limited to the immediate multi-omics context of an individual entity. Multiple entities can be analyzed together, potentially revealing non-trivial connections and these networks can be expanded to explore their functional neighborhood (for genes defined by co-expression, protein co-abundance or co-regulation networks and for metabolites by partial correlation networks). By providing additional filtering options, such as restricting co-expression links to specific brain regions or applying different significance cutoffs, users can create and explore highly context-specific networks that integrate results from various sources. Besides dynamic generation and interactive exploration of networks, we additionally interlink entities to external databases and provide downstream analysis tools, including the overlay of experimental data on differently expressed genes and differentially abundant proteins, as well as gene set and pathway enrichment analysis.

Building upon and extending previous efforts that have provided multi-omics integration solutions for multi-disease drug repositioning [100,101], genome-guided computational analysis of AD [102] and multi-omics annotation of individual targets [103], the AD Atlas provides users with an easy-to-use and fully network-based research platform. We have demonstrated the utility and flexibility of this resource to answer a wide array of research questions, from drug repositioning efforts to explorative analysis and contextualization of AD findings. For example, by investigating AD-associated genes, as well as those targeted by drug repositioning candidates proposed in the literature, we highlight the potential of the AD Atlas to guide computational drug repositioning

efforts. Using two established AD risk genes, *APOE* and *CLU*, we were able to identify multiple proposed repositioning candidates and show that, even without adding pathway information, extensive data on drug effects, or detailed information on effect directions, the AD Atlas is able to capture disease-relevant information and guide drug prioritization. Information on differentially expressed entities (DEG/DEP) can additionally be overlaid to explore the extent of dysregulation and prioritize drug repositioning candidates that are known to induce transcriptional effects opposing the observed direction in disease [38].

In addition to this hypothesis-driven approach to drug repositioning, we showed that the AD Atlas also allows the generation of data-driven disease networks that allow a more global and unbiased view on the disease. Furthermore, we explored the proposed benefits of statins for AD therapy and identified a link to neuroinflammation for a subset of compounds targeting *ITGAL*. Beyond applications related to drug discovery, the AD Atlas also provides a valuable platform to contextualize prior analysis results and hypotheses, by enabling researchers to easily validate and complement their findings with information from multiple heterogeneous resources and publications within the scope of one resource. For example, we extended the findings of a previous analysis that implicated three sphingomyelins in AD by adding additional layers of multi-omics data to gain further insights into the involvement of these metabolites in AD. In another analysis, we showcased how the AD Atlas can be used to inform novel hypotheses. We generated subnetworks surrounding marker genes for homeostatic microglia and disease-associated microglia to investigate the underlying molecular mechanisms of this transition and identify possible drug repositioning candidates. By focusing on genes that harbor disease-related as well as metabolite-associated SNPs and additionally are specific to the respective marker-gene networks, we found a potential link to blood androgens, suggesting a role of these metabolites in the transition of microglia to a disease-associated state in AD. Additional gene set enrichment analysis of the molecular network surrounding both marker genes was able to identify a promising drug repositioning candidate and

provided further insights into the underlying molecular mechanisms, suggesting a potential interplay between androgen signaling and Rho kinase activity in AD.

In summary, by providing an integrated global view on multi-omics results and by enabling tailored analyses through a user-friendly interface, the AD Atlas is a valuable resource for the validation and formulation of hypothesis, providing actionable insights that can be validated in follow-up analysis or experiments.

5 Limitations and outlook

The AD Atlas in its current form provides extended access to large-scale multi-omics data generated by the AMP-AD initiative, where single biological entities are embedded into their molecular multi-level context. As such, in its current version the resource has several limitations.

First, the underlying molecular framework is derived primarily in a data-driven way, which leads to a largely bias-free network representation of molecular data. Although the confidence in these data-driven representations will increase through inclusion of additional omics-studies, we neglect a large body of curated knowledge from experts and public databases on, for example, biochemical pathways or drug-targets. As we continue to extend the AD Atlas, integrating such data will be a valuable addition and provide a further layer of evidence.

Second, in the web interface we have so far only applied simplistic, minimal edge filtering criteria, i.e. using genome-wide, gene-wise or study-specific significance thresholds. While the underlying database contains potentially relevant evidence at even finer granularity (up to a raw p-value ≤ 0.05), this data is currently not accessible. Similarly, additional information, such as study sample size or number and consistency of effects reported in different studies are currently not compared or meta-analyzed. To this end, we aim to develop more sophisticated, context-dependent edge weighting and filtering criteria, as well as more comprehensive analysis reports.

Third, based on the currently limited availability of research data obtained in diverse populations/ethnicities, the data used to generate the AD Atlas are also not representative of these populations, which reduces the generalizability and transferability of the integrated molecular observations. However, the AMP-AD 2.0 program has the particular mission to expand research efforts to Latino and African American communities, and we will put a particular emphasis on integration of resulting findings from these diverse populations in the AD Atlas resource.

Lastly, we have so far focused on the development of tools for the exploration of hypothesis-guided local subnetworks, neglecting the potential of the network underlying the AD Atlas to identify disease modules globally in a more hypothesis-free manner. To this end, we are currently working on the implementation of established network analysis tools and machine learning technologies to extend and make available a more powerful analysis toolbox provided through the AD Atlas web interface in future releases.

DATA AVAILIBILITY

The AD Atlas is accessible via the user interface at adatlas.org. A comprehensive listing of the exact data sources that were integrated in the AD Atlas is given in the **Supplementary Material**. The results published here are in whole or in part based on data obtained from the AD Knowledge Portal (https://adknowledgeportal.org). Data used in the preparation of this article were further obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (https://adni.loni.usc.edu).

CONFLICTS OF INTEREST

P.M.D. has received research grants (through Duke University) from Lilly, Avanir, Bausch, Alzheimer's Drug Discovery Foundation. P.M.D. has received advisory fees from Verily, Otsuka, Genomind, Cogniciti, Clearview, VitaKey, Neuronix, Neuroglee, and Transposon Therapeutics. P.M.D. owns shares or options in UMethod, Evidation Health, Transposon, Marvel Biome, and Advera Health. P.M.D. serves on the board of Apollo. P.M.D. is a co-inventor (through Duke) on patents relating to dementia biomarkers, metabolomics, and therapies. K.N., A.J.S., R.K.D., G.K., and M.A. are co-inventors on patents on applications of metabolomics in Alzheimer's disease. A.J.S. has been supported by multiple grants from the National Institute on Aging (NIA), National Cancer Institute (NCI), and National Collegiatie Athletic Association (NCAA)/Department of Defense (DoD) as well as collaborative research support from Eli Lilly unrelated to the work reported here. He also received nonfinancial support from Avid Radiopharmaceuticals and Neurovision, and has served as a consultant to Arkley BioTek and Bayer. He received support from Springer-Nature as Editor in Chief of Brain Imaging and Behavior, outside the scope of the work submitted here. All other authors report no disclosures.

ACKNOWLEDGEMENTS

This work was supported by the National Institutes of Health/National Institute on Aging (NIA) through grants 1RF1AG057452, 1RF1AG058942, 1RF1AG059093, 1U01AG061359, 1U19AG063744,

and 1R01AG069901. K.S. was supported by the Biomedical Research Program at Weill Cornell Medicine in Qatar, a program funded by the Qatar Foundation. P.M.D. is supported by grants from the National Institutes of Health (NIH), the Cure Alzheimer's Fund, the Steve Aoki Foundation and the Karen L Wrenn Trust. A.K.G. is supported by the NIA (U24AG061340, NIA RF1AG057443). Data collection and sharing for this project was funded by the Alzheimer's Disease Metabolomics Consortium (National Institute on Aging R01AG046171, RF1AG051550 and 3U01AG024904-09S4), the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Funding sources had no role in study design, the collection, analysis and interpretation of data, the writing of the report, or in the decision to submit the article for publication.

AUTHOR CONTRIBUTIONS

Conceptualization: M.A.W., J.K., G.K., M.A.; data analysis and resource development: M.A.W., J.K., S.N., K.N., A.K.G., T.W., K.H., P.W., W.R.M., N.L., G.K., M.A.; funding acquisition: R.K.D., C.V.D., G.K., M.A.; Visualization: M.A.W.; supervision: R.K.D., G.K., M.A.; writing - original draft: M.A.W., S.N., P.M.D., G.K., M.A.; writing - review & editing: J.K., A.K.G., J.B., P.J.M., A.J.S., C.V.D., K.S.; all authors read and approved the final manuscript.

REFERENCES

[1] Long JM, Holtzman DM. Alzheimer disease: An update on pathobiology and treatment strategies. Cell 2019;179:312–39.

[2] Mullard A. Landmark Alzheimer's drug approval confounds research community. Nature 2021;594:309–10.

[3] Ferreira D, Wahlund L-O, Westman E. The heterogeneity within Alzheimer's disease. Aging 2018;10:3058–60.

[4] Sancesario GM, Bernardini S. Alzheimer's disease in the omics era. Clin Biochem 2018;59:9–16.

[5] Hodes RJ, Buckholtz N. Accelerating medicines partnership: Alzheimer's disease (AMP-AD) knowledge portal aids Alzheimer's drug discovery through open data sharing. Expert Opin Ther Targets 2016;20:389–91.

[6] Greenwood AK, Montgomery KS, Kauer N, Woo KH, Leanza ZJ, Poehlman WL, et al. The AD Knowledge Portal: A repository for multi-omic data on Alzheimer's disease and aging. Curr Protoc Hum Genet 2020;108:e105.

[7] Lysenko A, Roznovăţ IA, Saqi M, Mazein A, Rawlings CJ, Auffray C. Representing and querying disease networks using graph databases. BioData Mining 2016;9:23. https://doi.org/10.1186/s13040-016-0102-8.

[8] Lee B, Zhang S, Poleksic A, Xie L. Heterogeneous Multi-Layered Network Model for Omics Data Integration and Analysis. Front Genet 2020;10:1–11.

[9] Fang J, Pieper AA, Nussinov R, Lee G, Bekris L, Leverenz JB, et al. Harnessing endophenotypes and network medicine for Alzheimer's drug repurposing. Med Res Rev 2020;40:2386–426.

[10] Bellenguez C, Grenier-Boley B, Lambert J-C. Genetics of Alzheimer's disease: where we are, and where we are going. Curr Opin Neurobiol 2020;61:40–8.

[11] Wörheide MA, Krumsiek J, Kastenmüller G, Arnold M. Multi-omics integration in biomedical research – A metabolomics-centric review. Anal Chim Acta 2021;1141:144–62.

[12] Cunningham F, Achuthan P, Akanni W, Allen J, Amode MR, Armean IM, et al. Ensembl 2019. Nucleic Acids Res 2019;47:D745–51.

[13] Arnold M, Raffler J, Pfeufer A, Suhre K, Kastenmüller G. SNiPA: an interactive, genetic variant-centered annotation browser. Bioinformatics 2015;31:1334–6.

[14] GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. Science 2020;369:1318–30.

[15] Xie Z, Bailey A, Kuleshov MV, Clarke DJB, Evangelista JE, Jenkins SL, et al. Gene set knowledge discovery with Enrichr. Curr Protoc 2021;1:e90.

[16] Kolberg L, Raudvere U, Kuzmin I, Vilo J, Peterson H. gprofiler2 -- an R package for gene list functional enrichment analysis and namespace conversion toolset g:Profiler. F1000Res 2020;9:709.

[17] Alexa A, Rahnenfuhrer J. topGO: Enrichment Analysis for Gene Ontology. 2019.

[18] Jack CR, Bennett DA, Blennow K, Carrillo MC, Dunn B, Budd S, et al. NIA-AA Research Framework : Toward a biological definition of Alzheimer ' s disease. Alzheimers Dement 2018;14:535–62.

[19] Wan Y-W, Al-Ouran R, Mangleburg CG, Perumal TM, Lee TV, Allison K, et al. Meta-analysis of the Alzheimer's disease human brain transcriptome and functional dissection in mouse models. Cell Rep 2020;32:107908.

[20] Johnson ECB, Dammer EB, Duong DM, Ping L, Zhou M, Yin L, et al. Large-scale proteomic analysis of Alzheimer's disease brain and cerebrospinal fluid reveals early changes in energy metabolism associated with microglia and astrocyte activation. Nat Med 2020;26:769–80.

[21] Beecham GW, Hamilton K, Naj AC, Martin ER, Huentelman M, Myers AJ, et al. Genome-Wide Association Meta-analysis of Neuropathologic Features of Alzheimer's Disease and Related Dementias. PLoS Genet 2014;10. https://doi.org/10.1371/journal.pgen.1004606.

[22] 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 Neuropathol 2017;133:839–56.

[23] 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;45:1452–8.

[24] Marioni RE, Harris SE, Zhang Q, McRae AF, Hagenaars SP, Hill WD, et al. GWAS on family history of Alzheimer's disease. Transl Psychiatry 2018;8:0–6.

[25] 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 Aβ, tau, immunity and lipid processing. Nat Genet 2019;51:414–30.

[26] Suhre K, Shin SY, Petersen AK, Mohney RP, Meredith D, Wägele B, et al. Human metabolic individuality in biomedical and pharmaceutical research. Nature 2011;477:54–62.

[27] Shin S-Y, Fauman EB, Petersen A-K, Krumsiek J, Santos R, Huang J, et al. An atlas of genetic influences on human blood metabolites. Nat Genet 2014;46:543–50.

[28] Raffler J, Friedrich N, Arnold M, Kacprowski T, Rueedi R, Altmaier E, et al. Genome-Wide Association Study with Targeted and Non-targeted NMR Metabolomics Identifies 15 Novel Loci of Urinary Human Metabolic Individuality. PLoS Genet 2015;11. https://doi.org/10.1371/journal.pgen.1005487.

[29] Draisma HHM, Pool R, Kobl M, Jansen R, Petersen AK, Vaarhorst AAM, et al. Genome-wide association study identifies novel genetic variants contributing to variation in blood metabolite levels. Nat Commun 2015;6:21.

[30] Long T, Hicks M, Yu HC, Biggs WH, Kirkness EF, Menni C, et al. Whole-genome sequencing identifies common-to-rare variants associated with human blood metabolites. Nat Genet 2017;49:568–78.

[31] MahmoudianDehkordi S, Arnold M, Nho K, Ahmad S, Jia W, Xie G, et al. Altered bile acid profile associates with cognitive impairment in Alzheimer's disease—An emerging role for gut microbiome. Alzheimer's and Dementia 2019;15:76–92.

[32] Nho K, Kueider-Paisley A, MahmoudianDehkordi S, Arnold M, Risacher SL, Louie G, et al. Altered bile acid profile in mild cognitive impairment and Alzheimer's disease: Relationship to neuroimaging and CSF biomarkers. Alzheimer's and Dementia 2019;15:232–44.

[33] Arnold M, Nho K, Kueider-Paisley A, Massaro T, Huynh K, Brauner B, et al. Sex and APOE ε4 genotype modify the Alzheimer's disease serum metabolome. Nat Commun 2020;11:1148.

[34] Sieberts SK, Perumal TM, Carrasquillo MM, Allen M, Reddy JS, Hoffman GE, et al. Large eQTL meta-analysis reveals differing patterns between cerebral cortical and cerebellar brain regions. Sci Data 2020;7:340.

[35] Suhre K, Arnold M, Bhagwat AM, Cotton RJ, Engelke R, Raffler J, et al. Connecting genetic risk to disease end points through the human blood plasma proteome. Nat Commun 2017;8. https://doi.org/10.1038/ncomms14357.

[36] Krumsiek J, Suhre K, Evans AM, Mitchell MW, Mohney RP, Milburn MV, et al. Mining the Unknown: A Systems Approach to Metabolite Identification Combining Genetic and Metabolic Information. PLoS Genet 2012;8. https://doi.org/10.1371/journal.pgen.1003005.

[37] Cummings J, Feldman HH, Scheltens P. The "rights" of precision drug development for Alzheimer's disease. Alzheimers Res Ther 2019;11:1–14.

[38] Ballard C, Aarsland D, Cummings J, O'Brien J, Mills R, Molinuevo JL, et al. Drug repositioning and repurposing for Alzheimer disease. Nat Rev Neurol 2020. https://doi.org/10.1038/s41582-020-0397-4.

[39] Strittmatter WJ, Weisgraber KH, Huang DY, Dong LM, Salvesen GS, Pericak-Vance M, et al. Binding of human apolipoprotein E to synthetic amyloid beta peptide: isoform-specific effects and implications for late-onset Alzheimer disease. Proc Natl Acad Sci U S A 1993;90:8098–102.

[40] Lambert J-C, Heath S, Even G, Campion D, Sleegers K, Hiltunen M, et al. Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. Nat Genet 2009;41:1094–9.

[41] Harold D, Abraham R, Hollingworth P, Sims R, Gerrish A, Hamshere ML, et al. Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. Nat Genet 2009;41:1088–93.

[42] Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res 2016;44:W90-7.

[43] Gold M, Alderton C, Zvartau-Hind M, Egginton S, Saunders AM, Irizarry M, et al. Rosiglitazone monotherapy in mild-to-moderate Alzheimer's disease: results from a randomized, double-blind, placebo-controlled phase III study. Dement Geriatr Cogn Disord 2010;30:131–46.

[44] Harrington C, Sawchak S, Chiang C, Davies J, Donovan C, Saunders AM, et al. Rosiglitazone does not improve cognition or global function when used as adjunctive therapy to AChE inhibitors in mild-to-moderate Alzheimer's disease: two phase 3 studies. Curr Alzheimer Res 2011;8:592–606.

[45] Tariot PN, Schneider LS, Cummings J, Thomas RG, Raman R, Jakimovich LJ, et al. Chronic divalproex sodium to attenuate agitation and clinical progression of Alzheimer disease. Arch Gen Psychiatry 2011;68:853–61.

[46] Mdawar B, Ghossoub E, Khoury R. Selective serotonin reuptake inhibitors and Alzheimer's disease. Neural Regen Res 2020;15:41–6.

[47] Xie Y, Liu P-P, Lian Y-J, Liu H-B, Kang J-S. The effect of selective serotonin reuptake inhibitors on cognitive function in patients with Alzheimer's disease and vascular dementia: focusing on fluoxetine with long follow-up periods. Signal Transduct Target Ther 2019;4:30.

[48] Cummings J, Lee G, Ritter A, Sabbagh M, Zhong K. Alzheimer's disease drug development pipeline: 2020. Alzheimers Dement (N Y) 2020;6:e12050.

[49] Bauzon J, Lee G, Cummings J. Repurposed agents in the Alzheimer's disease drug development pipeline. Alzheimers Res Ther 2020;12:98.

[50] Hajjar I, Okafor M, McDaniel D, Obideen M, Dee E, Shokouhi M, et al. Effects of candesartan vs lisinopril on neurocognitive function in older adults with executive mild cognitive impairment: A randomized clinical trial. JAMA Netw Open 2020;3:e2012252.

[51] Hajjar I, Hart M, Chen Y-L, Mack W, Milberg W, Chui H, et al. Effect of antihypertensive therapy on cognitive function in early executive cognitive impairment: a double-blind randomized clinical trial. Arch Intern Med 2012;172:442–4.

[52] Rosenfeld CS, Shay DA, Vieira-Potter VJ. Cognitive effects of aromatase and possible role in memory disorders. Front Endocrinol (Lausanne) 2018;9:610.

[53] Decourt B, Lahiri DK, Sabbagh MN. Targeting tumor necrosis factor alpha for Alzheimer's disease. Curr Alzheimer Res 2017;14:412–25.

[54] Butchart J, Brook L, Hopkins V, Teeling J, Püntener U, Culliford D, et al. Etanercept in Alzheimer disease: A randomized, placebo-controlled, double-blind, phase 2 trial. Neurology 2015;84:2161–8.

[55] Torres-Acosta N, O'Keefe JH, O'Keefe EL, Isaacson R, Small G. Therapeutic potential of TNF-α inhibition for Alzheimer's disease prevention. J Alzheimers Dis 2020;78:619–26.

[56] De Sousa Rodrigues ME, Houser MC, Walker DI, Jones DP, Chang J, Barnum CJ, et al. Targeting soluble tumor necrosis factor as a potential intervention to lower risk for late-onset Alzheimer's

disease associated with obesity, metabolic syndrome, and type 2 diabetes. Alzheimers Res Ther 2019;12:1.

[57] MacPherson KP, Sompol P, Kannarkat GT, Chang J, Sniffen L, Wildner ME, et al. Peripheral administration of the soluble TNF inhibitor XPro1595 modifies brain immune cell profiles, decreases beta-amyloid plaque load, and rescues impaired long-term potentiation in 5xFAD mice. Neurobiol Dis 2017;102:81–95.

[58] Steele NZR, Carr JS, Bonham LW, Geier EG, Damotte V, Miller ZA, et al. Fine-mapping of the human leukocyte antigen locus as a risk factor for Alzheimer disease: A case-control study. PLoS Med 2017;14:e1002272.

[59] Sodero AO, Barrantes FJ. Pleiotropic effects of statins on brain cells. Biochim Biophys Acta Biomembr 2020;1862:183340.

[60] Liao JK, Laufs U. Pleiotropic effects of statins. Annu Rev Pharmacol Toxicol 2005;45:89–118.

[61] Haag MDM, Hofman A, Koudstaal PJ, Stricker BHC, Breteler MMB. Statins are associated with a reduced risk of Alzheimer disease regardless of lipophilicity. The Rotterdam Study. J Neurol Neurosurg Psychiatry 2009;80:13–7.

[62] Torrandell-Haro G, Branigan GL, Vitali F, Geifman N, Zissimopoulos JM, Brinton RD. Statin therapy and risk of Alzheimer's and age-related neurodegenerative diseases. Alzheimers Dement (N Y) 2020;6:e12108.

[63] Chu C-S, Tseng P-T, Stubbs B, Chen T-Y, Tang C-H, Li D-J, et al. Use of statins and the risk of dementia and mild cognitive impairment: A systematic review and meta-analysis. Sci Rep 2018;8. https://doi.org/10.1038/s41598-018-24248-8.

[64] Zissimopoulos JM, Barthold D, Brinton RD, Joyce G. Sex and race differences in the association between statin use and the incidence of Alzheimer disease. JAMA Neurol 2017;74:225–32.

[65] Geifman N, Brinton RD, Kennedy RE, Schneider LS, Butte AJ. Evidence for benefit of statins to modify cognitive decline and risk in Alzheimer's disease. Alzheimers Res Ther 2017;9. https://doi.org/10.1186/s13195-017-0237-y.

[66] Schultz BG, Patten DK, Berlau DJ. The role of statins in both cognitive impairment and protection against dementia: a tale of two mechanisms. Transl Neurodegener 2018;7. https://doi.org/10.1186/s40035-018-0110-3.

[67] 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:D1074–82.

[68] Zhang B, Gaiteri C, Bodea L-G, Wang Z, McElwee J, Podtelezhnikov AA, et al. Integrated systems approach identifies genetic nodes and networks in late-onset Alzheimer's disease. Cell 2013;153:707–20.

[69] Walling BL, Kim M. LFA-1 in T cell migration and differentiation. Front Immunol 2018;9:952.

[70] Naj AC, Jun G, Beecham GW, Wang L-S, Vardarajan BN, Buros J, et al. Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. Nat Genet 2011;43:436–41.

[71] Karch CM, Jeng AT, Nowotny P, Cady J, Cruchaga C, Goate AM. Expression of novel Alzheimer's disease risk genes in control and Alzheimer's disease brains. PLoS One 2012;7:e50976.

[72] Weitz-Schmidt G, Welzenbach K, Brinkmann V, Kamata T, Kallen J, Bruns C, et al. Statins selectively inhibit leukocyte function antigen-1 by binding to a novel regulatory integrin site. Nat Med 2001;7:687–92.

[73] Zenaro E, Pietronigro E, Della Bianca V, Piacentino G, Marongiu L, Budui S, et al. Neutrophils promote Alzheimer's disease-like pathology and cognitive decline via LFA-1 integrin. Nat Med 2015;21:880–6.

[74] Haure-Mirande J-V, Audrain M, Fanutza T, Kim SH, Klein WL, Glabe C, et al. Deficiency of TYROBP, an adapter protein for TREM2 and CR3 receptors, is neuroprotective in a mouse model of early Alzheimer's pathology. Acta Neuropathol 2017;134:769–88.

[75] Toledo JB, Arnold M, Kastenmüller G, Chang R, Baillie RA, Han X, et al. Metabolic network failures in Alzheimer's disease: A biochemical road map. Alzheimer's and Dementia 2017;13:965–84.

[76] Baloni P, Arnold M, Moreno H, Nho K, Buitrago L, Huynh K, et al. Multi-omic analyses characterize the ceramide/sphingomyelin pathway as a therapeutic target in Alzheimer's Disease. BioRxiv 2021. https://doi.org/10.1101/2021.07.16.21260601.

[77] Vasiliou S. Oral fingolimod for the treatment of relapsing-remitting multiple sclerosis. Drugs Today (Barc) 2010;46:315–25.

[78] Carreras I, Aytan N, Choi J-K, Tognoni CM, Kowall NW, Jenkins BG, et al. Dual dose-dependent effects of fingolimod in a mouse model of Alzheimer's disease. Sci Rep 2019;9:10972.

[79] Jęśko H, Wencel PL, Lukiw WJ, Strosznajder RP. Modulatory effects of fingolimod (FTY720) on the expression of sphingolipid metabolism-related genes in an animal model of Alzheimer's disease. Mol Neurobiol 2019;56:174–85.

[80] Ozakbas S, Piri Cinar B, Yigit P, Baba C, Sagici O, Multiple Sclerosis Research Group. Five-year real-world data on fingolimod treatment's effects on cognitive function. Mult Scler Relat Disord 2021;54:103089.

[81] 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;169:1276-1290.e17.

[82] Satoh J-I, Kino Y, Asahina N, Takitani M, Miyoshi J, Ishida T, et al. TMEM119 marks a subset of microglia in the human brain. Neuropathology 2016;36:39–49.

[83] Bennett ML, Bennett FC, Liddelow SA, Ajami B, Zamanian JL, Fernhoff NB, et al. New tools for studying microglia in the mouse and human CNS. Proc Natl Acad Sci U S A 2016;113:E1738-46.

[84] Ulland TK, Colonna M. TREM2 - a key player in microglial biology and Alzheimer disease. Nat Rev Neurol 2018;14:667–75.

[85] Dohle GR, Smit M, Weber RFA. Androgens and male fertility. World J Urol 2003;21:341–5.

[86] Davis SR, Wahlin-Jacobsen S. Testosterone in women--the clinical significance. Lancet Diabetes Endocrinol 2015;3:980–92.

[87] Nebel RA, Aggarwal NT, Barnes LL, Gallagher A, Goldstein JM, Kantarci K, et al. Understanding the impact of sex and gender in Alzheimer's disease: A call to action. Alzheimers Dement 2018;14:1171–83.

[88] Fuller SJ, Tan RS, Martins RN. Androgens in the etiology of Alzheimer's disease in aging men and possible therapeutic interventions. J Alzheimers Dis 2007;12:129–42.

[89] Rosario ER, Chang L, Head EH, Stanczyk FZ, Pike CJ. Brain levels of sex steroid hormones in men and women during normal aging and in Alzheimer's disease. Neurobiol Aging 2011;32:604–13.

[90] Franco-Bocanegra DK, McAuley C, Nicoll JAR, Boche D. Molecular mechanisms of microglial motility: Changes in ageing and Alzheimer's disease. Cells 2019;8:639.

[91] Novikova G, Kapoor M, Tcw J, Abud EM, Efthymiou AG, Chen SX, et al. Integration of Alzheimer's disease genetics and myeloid genomics identifies disease risk regulatory elements and genes. Nat Commun 2021;12:1610.

[92] Karch CM, Ezerskiy LA, Bertelsen S, Alzheimer's Disease Genetics Consortium (ADGC), Goate AM. Alzheimer's disease risk polymorphisms regulate gene expression in the ZCWPW1 and the CELF1 loci. PLoS One 2016;11:e0148717.

[93] Henderson BW, Gentry EG, Rush T, Troncoso JC, Thambisetty M, Montine TJ, et al. Rho-associated protein kinase 1 (ROCK1) is increased in Alzheimer's disease and ROCK1 depletion reduces amyloid-β levels in brain. J Neurochem 2016;138:525–31.

[94] Gentry EG, Henderson BW, Arrant AE, Gearing M, Feng Y, Riddle NC, et al. Rho kinase inhibition as a therapeutic for progressive supranuclear palsy and corticobasal degeneration. J Neurosci 2016;36:1316–23.

[95] Guo M-F, Zhang H-Y, Li Y-H, Gu Q-F, Wei W-Y, Wang Y-Y, et al. Fasudil inhibits the activation of microglia and astrocytes of transgenic Alzheimer's disease mice via the downregulation of TLR4/Myd88/NF-κB pathway. J Neuroimmunol 2020;346:577284.

[96] Chen J, Sun Z, Jin M, Tu Y, Wang S, Yang X, et al. Inhibition of AGEs/RAGE/Rho/ROCK pathway suppresses non-specific neuroinflammation by regulating BV2 microglial M1/M2 polarization through the NF-κB pathway. J Neuroimmunol 2017;305:108–14.

[97] Kroiss A, Vincent S, Decaussin-Petrucci M, Meugnier E, Viallet J, Ruffion A, et al. Androgen-regulated microRNA-135a decreases prostate cancer cell migration and invasion through downregulating ROCK1 and ROCK2. Oncogene 2015;34:2846–55.

[98] Roser A-E, Tönges L, Lingor P. Modulation of microglial activity by Rho-kinase (ROCK) inhibition as therapeutic strategy in Parkinson's disease and Amyotrophic lateral sclerosis. Front Aging Neurosci 2017;9:94.

[99] Horstman AM, Dillon EL, Urban RJ, Sheffield-Moore M. The role of androgens and estrogens on healthy aging and longevity. J Gerontol A Biol Sci Med Sci 2012;67:1140–52.

[100] Himmelstein DS, Baranzini SE. Heterogeneous Network Edge Prediction: A Data Integration Approach to Prioritize Disease-Associated Genes. PLoS Comput Biol 2015;11:1–27.

[101] Himmelstein DS, Lizee A, Hessler C, Brueggeman L, Chen SL, Hadley D, et al. Systematic integration of biomedical knowledge prioritizes drugs for repurposing. Elife 2017;6:1–35.

[102] Sügis E, Dauvillier J, Leontjeva A, Adler P, Hindie V, Moncion T, et al. HENA, heterogeneous network-based data set for Alzheimer's disease. Scientific Data 2019;6:1–18.

[103] 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;13:24.