

An integrative Multi-Omics Approach to Identify Cell-Type Specific Causal Genes and Druggable Targets for Alzheimer's Disease *in silico*

Student: Siqi Zhou; Supervisor: Yizhou Yu, L. Miguel Martins; Martin's Lab, MRC Toxicology Unit, University of Cambridge

Background

Alzheimer's disease (AD) is a complex neurodegenerative disorder afflicting 50 million people globally; patients survive an average of seven years post-diagnosis. However, no effective therapy to date exists for its cure. There is a pressing necessity to identify novel targets for effective therapies, a process that can be greatly accelerated by computational methods that access multiple layers of biological evidence and build on vast ready drug-gene interaction datasets.

Mendelian randomization (MR) is a powerful statistical method that establishes causal relationships between exposures and outcomes, by inferring causation instead of relevance by proxy to randomized controlled trials. Single-cell RNA sequencing (scRNA-seq) has revolutionized our understanding of cellular heterogeneity, providing an extensive yet detailed expression landscape at individual cell level. High-res sequencing unveils varying cell subpopulations and dynamic changes that underlie various biological processes and diseases, critical in drug targeting and administration.

This project combines previously generated MR results with transcriptomic analysis, and build upon newly generated GWAS, single cell RNA-seq, and proteomic data to identify cell-type specific causal genes with a causal contribution to Alzheimer's disease, highlights tissue specific genes at critical points of regulation in both transcription and translation, and identify those with druggable potential. Together, we find 51 known or novel locus of AD with 5 potential drug repurposing targets, and highlight the potential of several anti-cancer drugs on BRD2 in AD treatment. This computational pipeline will be taken to publication. *In vivo* tests of the aforementioned drugs in *Drosophila* AD models are under development as the next step.

Results

MR analysis. In our extensive MR screen, we used tissue-specific expression quantitative trait loci (eQTLs) from 109 tissues to potential causal relationships between gene expression and AD outcome. The preliminary screen pinpointed brain and blood tissues with the most significant enrichment for causal exposures. A subsequent screen incorporating single-cell resolution QTL data yielded 68 brain-specific and 146 blood-specific AD-associated exposures ($p < 5 \times 10^{-8}$), implicating 45 and 88 genes in distinct cell types in the brain and blood.

scRNA-seq. To validate single-cell eQTL results, regression tests between clinical AD status and genetic variants were conducted in two separate brain and blood scRNA-seq datasets (Figure 1A,B). In brain post-mortem section represented by dorsal-lateral prefrontal cortex (DLPFC), 12 protein-coding genes and one diverse transcript showed significance in AD status and progression (Figure 1C,D). The peripheral blood mononuclear cells (PBMC) highlight a richer set of 38 transcripts from 36 protein genes and 2 antisense lncRNA, which will be discussed further in the discussion section (Figure 1E). In DLPFC, astrocytes showed notable changes, and CD4 and CD8 lymphocytes were predominantly altered among PBMC.

Pathway analysis. Protein-Protein Interaction (PPI) networks of the two gene sets were constructed using the STRING database (Figure 2). Both scRNA-seq datasets indicated a heightened enrichment in gene interaction (DLPFC, $p = 6.48 \times 10^{-5}$; PBMC, $p = 0.00143$), with most edges connected by co-expression or genetic neighbourhood; both show centres of enrichment around the chr17:q21.31 locus. We also identified multiple proteins in specific biological processes like the MHC II and NSL transcription complex.

Drug identification. Quality control measures were employed to ensure consistency in drug issuing. 26 genes were selected for consistent direction of effect across 1) cell types, 2) tissue types, 3) MR meta-analysis methods. High-potency functionally interactive or binding drug-gene interactions were identified based on the ChEMBL Database (Szkarczyk et al., 2015), and filtered referring to methods provided by the python drug2cell package (Kanemaru et al., 2023). In the prioritized gene list, 21 drug-gene interactions indicate 6 genes (MAPT, CRHR1, BRD2, RPS15, MARK4, PHKG2) with reported drug-gene interaction where drugs have entered clinical trials, as detailed in Table 1.

Proteomic integration. We cross-referenced transcriptome results with available proteome expression data (Haytural et al., 2021). Despite the statistically significant RNA-level changes in 17 genes, only 6 exhibited alterations at the protein level, suggesting the potential of RNAi therapeutics. Interestingly, genes including MAPT present divergences between RNA and protein changes during AD progression, which implicate extra regulatory layers during AD pathogenesis.

Discussion

MR and scRNA-seq results. The three predominant hypotheses for AD include the tau hypothesis, the amyloid cascade hypothesis, and the less understood autoimmune hypothesis, suggesting that AD can be initiated by Blood Brain Barrier (BBB) dysfunction and leukocyte infiltration (Montagne et al., 2017; Arshavsky, 2020). We find key AD-causal genes in the immune response pathway, including HLA-DQA1/2 and PICALM, and several other immune response genes including BIN1, FCER1G, BCKDK, and QPCTL at both tissue and single-cell resolutions, which support the hypothesis that an immune response promoting leukocyte proliferation and migration during BBB dysfunction may initiate amyloid-beta and tau turnover as well as neuroinflammation.

We report several novel SNPs that significantly increase AD risk (reporting Odds Ratio as high as 6.23×10^{13}) which will be further discussed in the paper to be published.

PPI. We reiterate the significance of the established genetic locus chr17:q21.31, especially KANSL1, in AD etiology. We also report TOMM40, located adjacent to the CSF biomarker APOE on chromosome 19, as causal, both of which are closely related to mitochondrial dynamics; TOMM40 moreover shows potential interaction with immune-response gene PICALM (co-expression score=0.042, from GEO microarray and String). This, combined with the fact that the three AD hypotheses are not mutually exclusive, lead us to believe that higher level dysregulation may occur prior and leading to the immune response or tau cascade, such as within the fractal globule.

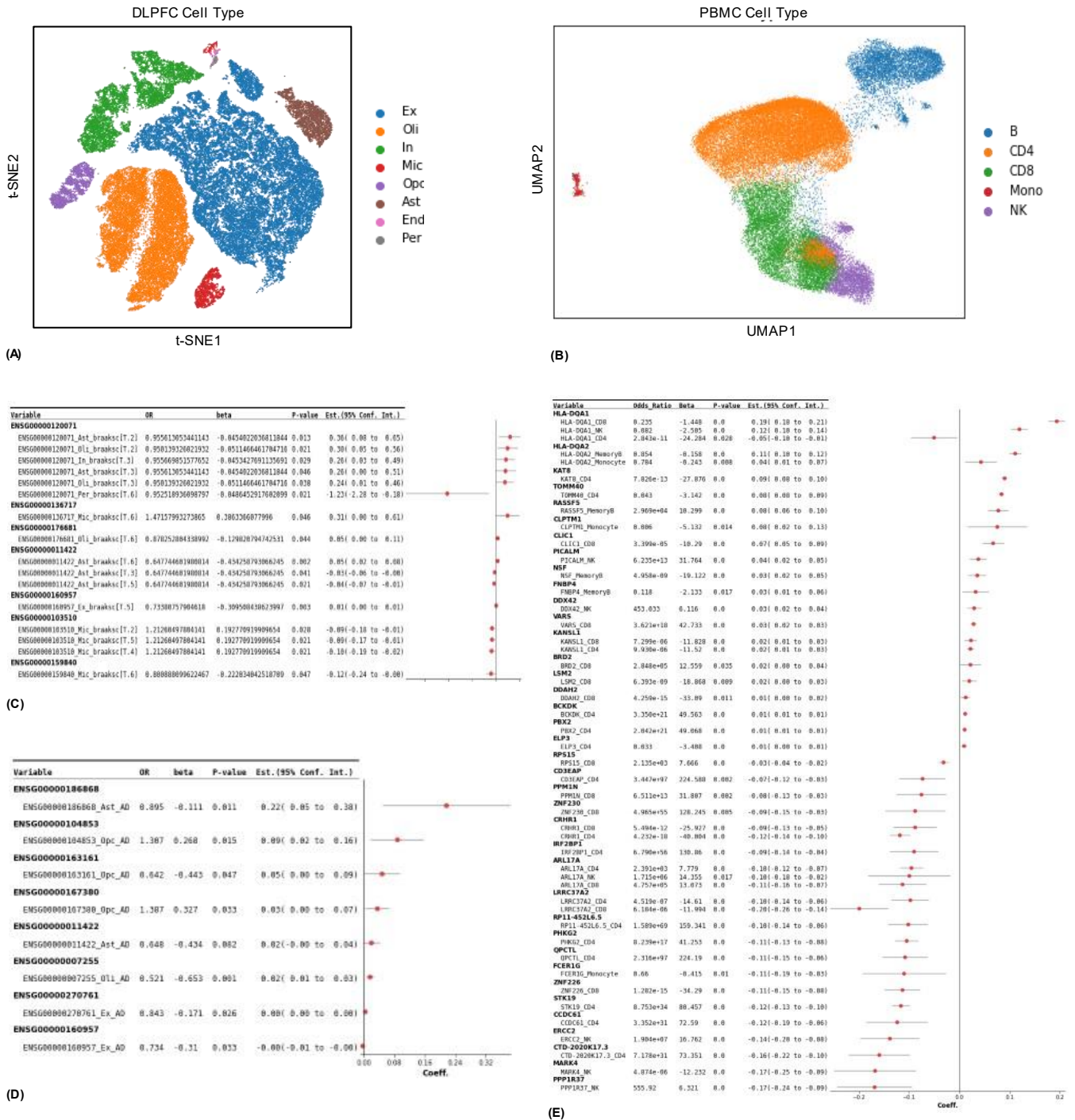
Indeed, chromatin dysregulation is evident with our findings on causal downregulation of ERCC2 and ERCC3 in the DNA repair factor XPD, co-expressing subunits CD3EAP of RNAP I, and DNA helicase Q4 subunit RECQL4, indicating compromised nuclear excision repair mechanism in AD patients. This is complemented by the causal roles of BRD2 and CTD-2020K17.3 (antisense to FMNL1) in chromosomal interactions, where FMNL1 is recently discovered as a tissue-specific non-expressing distal regulatory element to MAPT and MAP3K14 (Rogers et al., 2023).

Druggability. Amongst 6 genes that encode for efficacy targets of approved or clinical-phase drugs, BRD2, in line with our chromatin dysfunction prediction, shows high potential for drug repurposing. As a DNA remodeling factor that protects against H4 acetylation, our analysis suggests BRD2 exposures increase AD risk (OR=12.559, beta=2.84e6, expression fold change=0.02). Its expression can be inhibited by potent small molecules, including but not limited to Birabresib, Molibresib, and RO-6870810 to treat hematologic solid tumor and MYC dysregulation (Lewin et al., 2018; Cousin et al., 2022; Piha-Paul et al., 2019; Shapiro et al., 2021; Dickinson et al., 2021], and Apabetalone aimed at treating patients with diabetes and acute coronary syndrome (Nicholls et al., 2021). Another small molecule, JQ1, already shows potential in improving cognitive performance in APP mice models (Benito et al., 2017).

Limitations and future plans. Despite the effort to use the largest publicly available datasets, the power of statistical tests can be limited, especially by the number of individuals involved, as such with PBMC scRNA-seq which is severely lacking in large-scale studies compared to that of the AD post-mortem brain. The difference in significance in proteome and transcriptome expression also suggests the need for enriching proteomic data.

Having established the potential of BRD2, we plan to verify the effect of identified drugs in the *Drosophila* or mice ADD models. Both behavioral and molecular measurements would be conducted to measure if identified drugs can extend lifespan, mitigate memory loss, improve behavioral patterns, and restore cellular conditions including mitochondria potential. Taken together, the *in silico* MR and scRNA modeling and *in vivo* experiments would form a pipeline for developing and repurposing drugs for neurodegeneration.

Figures and Tables



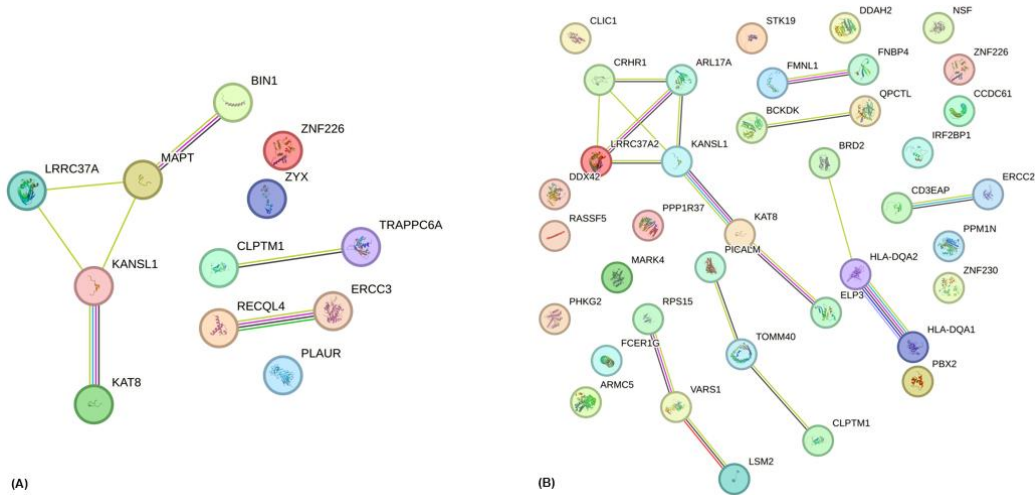


Figure 2. Protein-Protein Interaction Network as calculated by STRING database. Cyan: known interaction from curated databases; magenta: experimentally determined interactions; green: gene neighborhood; red: gene fusions; blue: co-occurrence; light green: textmining; black: co-expression; light purple: protein homology. **A)** 16 protein coding genes and CD2AP corresponding to identified CD2AP-DT in DLPFC. **B)** 36 protein coding genes and 2 proteins corresponding to their identified antisense RNA.

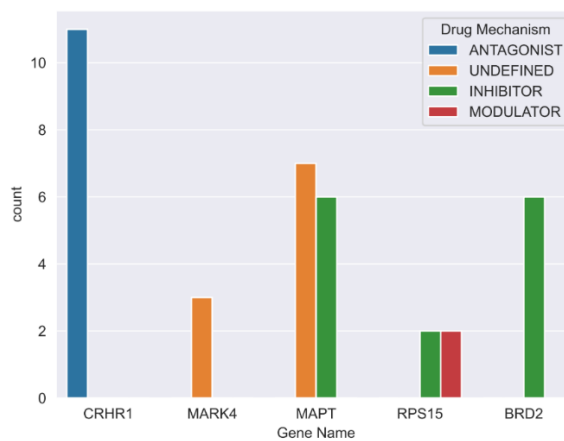


Figure 3. Count of drugs which are a) interacting with one of 26 potential drug targets and b) have a known direction of effect. Drug-gene interactions with known affinity but unknown effect are not reported in this graph.

activitiesla	activitiesle	activitieslr	activitiesls	activitieslt	activitiesls	activitiesls	activitiesls	activitiesls	activitiesls	assayslsides	assayslssay_type	assaysltid	assayslconfidence_score	assayslschembi_id	drugmolregno	drug_medi	drug_mechanis m
420088	118904	51112	102752	6.62	Ki	=	240	nM	Ki	Binding af B	124	8	CHEMBL664958	102752	Corticotroj	ANTAGONIST	
3433569	2381895	494923	515198	7.18	pIC50	=	66.07	nM	IC50	Displacem B	124	9	CHEMBL1004666	515198	Corticotroj	ANTAGONIST	
3433570	2381932	494924	515198	7.14	pA2	=	72.44	nM	Kd	Antagonis F	124	9	CHEMBL1004667	515198	Corticotroj	ANTAGONIST	
3433571	2381933	494942	515198		Activity				Activity	Not Active Agonist acF	124	9	CHEMBL1008014	515198	Corticotroj	ANTAGONIST	
3593875	2528492	501842	520543	8.35	IC50	=	4.5	nM	IC50	Displacem B	124	9	CHEMBL994499	520543	Corticotroj	ANTAGONIST	
3593876	15455721	1482271	520543	8.21	IC50	=	6.1	nM	IC50	Antagonis B	124	9	CHEMBL3539309	520543	Corticotroj	ANTAGONIST	
3593879	18000968	1653073	520543		Ratio IC50 >		29		Ratio IC50	Ratio of IC B	124	9	CHEMBL4002328	520543	Corticotroj	ANTAGONIST	
3919199	9516721	809161	494834	7	pKi	=	100	nM	Ki	active	PUBCHEM F	101317	8	CHEMBL1963713			
4568021	3604367	685890	714326	7	pIC50	=	100	nM	IC50	Antagonis F	124	9	CHEMBL1292646	714326	Corticotroj	ANTAGONIST	
4568022	3604375	685890	714326	8.2	pKi	=	6.31	nM	Ki	Antagonis F	124	9	CHEMBL1292646	714326	Corticotroj	ANTAGONIST	
5130506	4516677	688171	94571	7.6	Potency	=	25.1	nM	Potency	Inconclusi	PUBCHEM F	103657	9	CHEMBL1614250			
5130507	4516677	688171	94571	7.6	Potency	=	25.1	nM	Potency	Inconclusi	PUBCHEM F	103657	9	CHEMBL1614250			
5130508	4516677	688171	94571	7.6	Potency	=	25.1	nM	Potency	Inconclusi	PUBCHEM F	103657	9	CHEMBL1614250			
5890044	4527053	688171	373953	7.6	Potency	=	25.1	nM	Potency	Inconclusi	PUBCHEM F	103657	9	CHEMBL1614250			
5890045	4527053	688171	373953	7.6	Potency	=	25.1	nM	Potency	Inconclusi	PUBCHEM F	103657	9	CHEMBL1614250			
5890046	4527053	688171	373953	7.6	Potency	=	25.1	nM	Potency	Inconclusi	PUBCHEM F	103657	9	CHEMBL1614250			
5890047	4527053	688171	373953	7.6	Potency	=	25.1	nM	Potency	Inconclusi	PUBCHEM F	103657	9	CHEMBL1614250			
7456587	6335363	764455	1171281	8.4	IC50	=	4	nM	IC50	Displacem B	124	9	CHEMBL1821007	1171281	Corticotroj	ANTAGONIST	
7456588	6335393	764454	1171281	7.4	EC50	=	40	nM	EC50	Antagonis F	124	9	CHEMBL1821006	1171281	Corticotroj	ANTAGONIST	
8695475	9516718	809161	460144	7	pKi	=	100	nM	Ki	active	PUBCHEM F	101317	8	CHEMBL1963713			
10607977	9516698	809161	630975	7.3	pKi	=	50.12	nM	Ki	active	PUBCHEM F	101317	8	CHEMBL1963713			
22433509																	
22434152															1381390	80S Ribosc	INHIBITOR
22434156															1947706	Bromodoin	INHIBITOR
22434165															694811	Bromodoin	INHIBITOR
22434169															2335429	Bromodoin	INHIBITOR
22434173															1566719	Bromodoin	INHIBITOR
22435038															2335394	Bromodoin	INHIBITOR
22435077															2341275	Bromodoin	INHIBITOR
22435111															2335831	Microtubuj	INHIBITOR
22435259															2335760	80S Ribosc	INHIBITOR
22435340															426110	80S Ribosc	MODULATOR
22435963															2335715	80S Ribosc	MODULATOR
22435964															2336064	Microtubuj	INHIBITOR
22435965															2197913	Microtubuj	INHIBITOR
22437968															2335992	Microtubuj	INHIBITOR
22438678															2486662	Microtubuj	INHIBITOR
															2465043	Microtubuj	INHIBITOR

Methods

MR. Bulk tissue RNA-seq were downloaded from the EMBL-EBI eQTL catalogue on 10 Nov 2022 (Kerimov et al., 2021), single-cell brain eQTL summary statistics ($P < 0.05$) on 5 Jan 2023 (Bryois et al., 2019). Summary statistics of SNPs associated with AD for the outcome dataset were obtained from the largest-to-date GWAS study (Bellenguez et al.; 2022).

scRNA-seq processing. The 48 sample prefrontal cortex scRNA-seq was curated by Mathys et al. (Mathys et al., 2019), and fitted to a linear mixed model using python *statsmodels*; cell clusters was as provided by the paper. For blood samples a less extensive 5 sample PBMC RNA-seq (Xu & Jia, 2021) was selected. Cells were reclustered using the scanpy package for poor overlap of patients in the original clustering; more specifically, one round of rough cell type identification by tsne (PCs=10), leiden, and marker genes in scanpy rank_genes_groups as provided by the paper, and a second round of identification using ingest and bbknn with patient GSM5494110's projected to PBMC 3k as reference. Group 22 was excluded due to PF4 and PPBP expression suggesting mixture with platelets. The dataset was fitted by regression.

ChEMBL and drug2cell. The ChEMBL database was used to identify drug-gene interactions; selection of interaction threshold and assay type followed those provided in the drug2cell package. We further selected for compounds with potency concentration below 100nM.

Acknowledgements

I would like to extend my deepest gratitude to my supervisor, Yizhou Zhou, and PI, L. Miguel Martins, for their unwavering support throughout the duration of this project, for the myriad of opportunities they provided, from involving me in lab dailies and meetings to granting me the chance to attend various biomedical and biochemical conferences. I would like to express my sincere thanks to Bryan Tan, who meticulously curated the MR analysis data that were the cornerstone of this project. Their work, advice, and support have been instrumental in the formation of this report.

References

- Arshavsky YI. Alzheimer's Disease: From Amyloid to Autoimmune Hypothesis. *Neurosci.* 2020;26: 455–470. doi:10.1177/1073858420908189
- Bellenguez C, Küçükali F, Jansen IE, Kleindam L, Moreno-Grau S, Amin N, et al. New insights into the genetic etiology of Alzheimer's disease and related dementias. *Nat Genet.* 2022;54: 412–436. doi:10.1038/s41588-022-01024-z
- Benito E, Ramachandran B, Schroeder H, Schmidt G, Urbanke H, Burkhardt S, et al. The BET/BRD inhibitor JQ1 improves brain plasticity in WT and APP mice. *Transl Psychiatry.* 2017 Sep;7(9):e1239.
- Bryois J, Calini D, Macnair W, Foo L, Urich E, Ortmann W, et al. Cell-type-specific cis-eQTLs in eight human brain cell types identify novel risk genes for psychiatric and neurological disorders. *Nat Neurosci.* 2022;25: 1104–1112. doi:10.1038/s41593-022-01128-z
- Cousin S, Blay JY, Garcia IB, de Bono JS, Le Tourneau C, Moreno V, et al. Safety, pharmacokinetic, pharmacodynamic and clinical activity of molibresib for the treatment of nuclear protein in testis carcinoma and other cancers: Results of a Phase I/II open-label, dose escalation study. *Int J Cancer.* 2022 Mar 15;150(6):993–1006.

Dickinson M, Briones J, Herrera AF, González-Barca E, Ghosh N, Cordoba R, et al. Phase 1b study of the BET protein inhibitor RO6870810 with venetoclax and rituximab in patients with diffuse large B-cell lymphoma. *Blood Adv.* 2021 Nov 22;5(22):4762–70.

Esteras, Noemí, Carolina Alquézar, and Ana de la Encarnación. "Lymphocytes in Alzheimer's disease pathology: Altered signaling pathways." *Current Alzheimer Research* 13.4 (2016): 439-449.

Giannakopoulos, Panteleimon, et al. "Cerebral cortex pathology in aging and Alzheimer's disease: a quantitative survey of large hospital-based geriatric and psychiatric cohorts." *Brain Research Reviews* 25.2 (1997): 217-245.

Haytural H, Benfeitas R, Schedin-Weiss S, Bereczki E, Rezeli M, Unwin RD, et al. Insights into the changes in the proteome of Alzheimer disease elucidated by a meta-analysis. *Sci Data.* 2021 Dec 3;8:312.

Hyman, Bradley T., et al. "Alzheimer's disease: cell-specific pathology isolates the hippocampal formation." *Science* 225.4667 (1984): 1168-1170.

Kanemaru K, Cranley J, Muraro D, Miranda AMA, Ho SY, Wilbrey-Clark A, et al. Spatially resolved multiomics of human cardiac niches. *Nature.* 2023 Jul;619(7971):801–10.

Kaye, Jeffrey A., et al. "Volume loss of the hippocampus and temporal lobe in healthy elderly persons destined to develop dementia." *Neurology* 48.5 (1997): 1297-1304.

Kerimov N, Hayhurst JD, Peikova K, Manning JR, Walter P, Kolberg L, et al. A compendium of uniformly processed human gene expression and splicing quantitative trait loci. *Nat Genet.* 2021;53: 1290–1299. doi:10.1038/s41588-021-00924-w

Leuner, Kristina, et al. "Peripheral mitochondrial dysfunction in Alzheimer's disease: focus on lymphocytes." *Molecular neurobiology* 46 (2012): 194-204.

Lewin J, Soria JC, Stathis A, Delord JP, Peters S, Awada A, et al. Phase 1b Trial With Birabresib, a Small-Molecule Inhibitor of Bromodomain and Extraterminal Proteins, in Patients With Selected Advanced Solid Tumors. *J Clin Oncol.* 2018 Oct 20;36(30):3007–14.

Nicholls SJ, Schwartz GG, Buhr KA, Ginsberg HN, Johansson JO, Kalantar-Zadeh K, et al. Apabetalone and hospitalization for heart failure in patients following an acute coronary syndrome: a prespecified analysis of the BETonMACE study. *Cardiovasc Diabetol.* 2021 Jan 7;20:13.

Mahapatra, Gargi, et al. "Blood-based bioenergetic profiling reveals differences in mitochondrial function associated with cognitive performance and Alzheimer's disease." *Alzheimer's & Dementia* 19.4 (2023): 1466-1478.

Mathys H, Davila-Velderrain J, Peng Z, Gao F, Mohammadi S, Young JZ, et al. Single-cell transcriptomic analysis of Alzheimer's disease. *Nature.* 2019 Jun;570(7761):332–7.

Montagne A, Zhao Z, Zlokovic B V. Alzheimer's disease: A matter of blood–brain barrier dysfunction? *J Exp Med.* 2017;214: 3151–3169. doi:10.1084/jem.20171406

Piha-Paul SA, Hann CL, French CA, Cousin S, Braña I, Cassier PA, et al. Phase 1 Study of Molibresib (GSK525762), a Bromodomain and Extra-Terminal Domain Protein Inhibitor, in NUT Carcinoma and Other Solid Tumors. *JNCI Cancer*

Spectr. 2019 Nov 6;4(2):pkz093.

Rogers BB, Anderson AG, Lauzon SN, Davis MN, Hauser RM, Roberts SC, et al. MAPT expression is mediated by long-range interactions with cis-regulatory elements. *bioRxiv*. 2023 Apr 11;2023.03.07.531520.

Shapiro GI, LoRusso P, Dowlati A, T Do K, Jacobson CA, Vaishampayan U, et al. A Phase 1 study of RO6870810, a novel bromodomain and extra-terminal protein inhibitor, in patients with NUT carcinoma, other solid tumours, or diffuse large B-cell lymphoma. *Br J Cancer*. 2021 Feb;124(4):744–53.

Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res*. 2015 Jan;43(Database issue):D447-452.

Tan, Z. S., et al. "Inflammatory markers and the risk of Alzheimer disease: the Framingham Study." *Neurology* 68.22 (2007): 1902-1908.

Xu H, Jia J. Single-Cell RNA Sequencing of Peripheral Blood Reveals Immune Cell Signatures in Alzheimer's Disease. *Front Immunol*. 2021 Aug 9;12:645666.