Supplemental Methods

snRNA-seq data quality control, dimensionality reduction and clustering

The snRNA-seq count matrices were processed using the *Scanpy* (v1.9.5) package(Wolf et al. 2018) in a Python 3.10 environment. Quality control metrics were initially computed using the `calculate_qc_metrics` function. Genes expressed in fewer than 10 cells and cells with fewer than 200 detected genes were excluded of the analysis. Outliers for `total_counts` and `n_genes_by_counts` were identified using the interquartile range (IQR) method. First, the 75th percentile (Q3) and 25th percentile (Q1) of the respective fields were calculated. The IQR was then determined as: $IQR = Q3 - Q1$. Based on this, lower and upper thresholds were defined as: $Lower\ threshold = Q1 - 3 \times IQR$; $Upper\ threshold = Q3 + 3 \times IQR$. Values outside this range were flagged as potential outliers, and only cells with values falling within the lower and upper thresholds were retained. For mitochondrial (`pct_counts_mt`) and ribosomal (`pct_counts_rb`) percentages, outlier thresholds were determined by combining the 3×IQR rule with fixed cutoffs. The upper threshold was computed using the 3×IQR rule: Upper threshold (IQR) = Q3 + 3 × IQR. This threshold was then compared with fixed cutoffs of 10% for `pct_counts_mt` and 20% for `pct_counts_rb`. The final upper threshold was the maximum of the IQR-based threshold and the fixed cutoff. Cells with values exceeding this final upper threshold were classified as outliers and excluded from further analysis. Finally, doublets were identified and excluded using the *Scrublet* tool (Wolock et al. 2019), further refining the dataset for downstream analyses.

We normalized the processed count matrices using a shifted logarithmic transformation to stabilize variance(Heumos et al. 2023). Principal components (PCs)

were computed to reduce dimensionality, and the data were clustered using the Leiden algorithm(Heumos et al. 2023; Traag et al. 2019). To visualize the clustering results, we utilized the Uniform Manifold Approximation and Projection (UMAP) technique(McInnes et al. 2020). Major cell type annotations for each cluster were assigned based on established marker genes(Mathys et al. 2023; Xiong et al. 2023; Gabitto et al. 2024; Morabito et al. 2021; Emani et al. 2024) (**Supporting Information Fig. S1**). We also compared these annotations with those reported in the original paper and found them to be consistent(Gabitto et al. 2024). Neuron cell subtypes were labeled according to the classifications in the original paper(Gabitto et al. 2024).

Differential expression gene (DEG) and gene functional enrichment analysis

We employed the R package *muscat* to identify differentially expressed genes (DEGs) by aggregating single-cell data into pseudobulk profiles, while adjusting for sex, APOE4 status, and age as covariates(Crowell et al. 2020). Each gene was tested for expression changes within individual clusters, resulting in a total of "#genes × #clusters" tests for each comparison of interest(Crowell et al. 2020). Specifically, we focused on condition-related changes across AD progression stages, analyzing six comparison groups: Low-AD versus (vs.) Non-AD, Inter-AD vs. Non-AD, High-AD vs. Non-AD, Inter-AD vs. Low-AD, High-AD vs. Low-AD, High-AD vs. Inter-AD. DEGs were defined as genes meeting the criteria of an adjusted p-value < 0.05 and an absolute log fold change (|logFC|) > 0.5, ensuring statistical significance and biologically meaningful changes.

Gene functional enrichment analysis was conducted using the *clusterProfiler* package (v4.10.1)(Yu et al. 2012; Wu et al. 2021; Xu et al. 2024b). Gene Ontology (GO) enrichment was evaluated for terms related to Biological Process (BP) and Molecular

Function (MF) across different stages of AD progression for each cell type. Enrichment results were filtered using the Benjamini-Hochberg (BH) method (adjusted *p*-value < 0.05).

snATAC-seq data quality control, dimensionality reduction and clustering

The snATAC-seq fragment data were processed using the *ArchR* R package (v1.0.3)(Granja et al. 2021) in an R 4.4.2 environment, employing the hg38 reference genome for annotation. For initial quality control, three key metrics were assessed: the number of unique nuclear fragments, the signal-to-background ratio calculated by transcription start site (TSS) enrichment score, and the fragment size distribution(Granja et al. 2021). To ensure sufficient data for reliable analysis, cells with fewer than 1,000 fragments were excluded (`minFrags = 1000`). The signal-to-background ratio was quantified using TSS enrichment score, with a minimum threshold of 4 (`minTSS = 4`). Cells not meeting these criteria were filtered out. To further enhance data quality, doublets were identified and removed using the *ArchR* functions *`addDoubletScores`* and *`filterDoublets`*, with a filter ratio of 1.

Dimensionality reduction was performed using the iterative Latent Semantic Indexing (LSI) approach implemented in the *ArchR* package via the *`addIterativeLSI`* function(Granja et al. 2021). The process utilized three iterations (`iterations = 3`), with the final iteration sampling 30,000 cells (`sampleCellsFinal = 30,000`). All other parameters are set to their default values. Clustering was performed using the *`addClusters`* function in *ArchR*, initially with a resolution of 1 and finalized with a resolution of 0.4. Clusters containing fewer than 100 cells were excluded as low-quality clusters. Marker genes for each cluster were

identified using the `getMarkerFeatures` function, which uses the Gene Activity Score (accessibility around each gene) as a proxy for gene expression(Ober-Reynolds et al. 2023). Cell types were annotated based on established marker genes (**Supporting Information Fig. S1**). Gene activity scores were visualized by UMAP, with smoothing applied using the MAGIC algorithm(van Dijk et al. 2018) to enhance visualization.

Peak calling and annotation

Peak calling was performed using *MACS2*(Zhang et al. 2008) with default parameters. Cell type-specific marker peaks were identified using the `getMarkerFeatures` function in the *ArchR* package. CREs specific to each cell type were further characterized by integrating these marker peaks with annotations from the hg38 reference genome. Transcription factor (TF) motif enrichment within these marker peaks was analyzed using the cisBP database(Weirauch et al. 2014) through the `addMotifAnnotations` and `peakAnnoEnrichment` functions, specifying `peakAnnotation = 'Motif'`. Additionally, footprinting analysis for motifs of interest was conducted using the `getFootprints` function in *ArchR*(Granja et al. 2021), providing insights into motif accessibility and regulatory activity.

Integration of snRNA- and snATAC -seq datasets

We integrated snATAC-seq data with corresponding snRNA-seq data for each stage of AD progression, following previously established methods(Ober-Reynolds et al. 2023). This integration was performed using the `addGeneIntegrationMatrix` function in the *ArchR* package, which leverages Seurat's `FindTransferAnchors` function to align datasets through canonical correlation analysis (CCA). For the integration process, we specified `nGenes = 2,000` and `dimsToUse = 1:40`. Initially, we conducted

unconstrained integration by setting `addToArrow = FALSE` to create a preliminary alignment. We subsequently refined the integration results by applying a constrained approach.

Linked AD associated GWAS loci with cCRE

We obtained GWAS summary statistics from the study by Bellenguez et al(Bellenguez et al. 2022) to identify AD-associated loci. Significant AD-associated GWAS loci were identified by applying a *p*-value threshold of *p*-value < 5e8. Additionally, We got AD lead SNPs of this study(Bellenguez et al. 2022)
from https://www.ebi.ac.uk/gwas/studies/GCST90027158, which excludes 23andMe samples. In total, we identified 88 lead SNPs. We also identified cell type-specific CREs from the significant cell marker peaks (FDR ≤ 0.05 and $\log_2$FC > 0.5). To establish functional links between these AD-associated loci and cCREs, we determined overlaps between the genomic positions of the AD-associated loci and CREs.

1.1. Identification of positive TF-regulators and putative targets

TFs can be identified based on changes in chromatin accessibility at DNA-binding motif sites using ATAC-seq. However, precise identification is often hindered due to the similarities of binding motifs among certain TF families by position weight matrices (PWMs)(Granja et al. 2021). To address this, we integrated gene expression data to identify TFs with expression that correlated positively with chromatin accessibility changes using *ArchR* package as described previously(Granja et al. 2021; Ober-Reynolds et al. 2023). This method correlates chromVAR deviation z-scores of TF motifs with the gene expression of corresponding TF across low-overlapping cell aggregates. Initially, deviant TF motifs were identified by calculating the maximum delta

in deviation z-scores across clusters, enabling the stratification of motifs based on inter-cluster variability. We then employed the `correlateMatrices` function in *ArchR* to compute the correlation between the gene expression matrix and the motif matrix. These correlations are assessed across many low-overlapping cell aggregates identified in the lower dimension space(Granja et al. 2021). Positive TF regulators were defined based on the following stringent criteria: a correlation coefficient (cor) between motif activity and gene expression greater than 0.5, an adjusted *p*-value below 0.01, and a maximum inter-cluster z-score variation exceeding 0.75.

We identified the regulatory targets of these positive TF-regulators following the methodology described in the previous study(Ober-Reynolds et al. 2023). First, we calculated the Pearson correlation coefficient between the chromVAR motif activity of positive TF regulators and the integrated expression levels of all expressed genes. We then computed the linkage score (LS) for each gene-TF pair using the following equation:

$$LS_g = \sum_{k=1}^{n} R_k^2 \, MS_k \qquad (1)$$

Here, LS_g: Linkage score for gene g. n: Number of peaks linked to gene g. R_k: Pearson correlation coefficient between peak k and gene g. MS_k: Motif score for the TF motif present in peak k. This scoring method ensured that higher linkage scores indicate stronger regulatory relationships, reflecting a greater number of linked peaks containing the TF motif, as well as strongly correlated peaks enriched for the motif.

Putative target genes for positive TF-regulators were filtered based on an LS threshold set at the 80th percentile of all LS values, retaining genes with LS values exceeding this threshold. Additionally, only genes with a correlation coefficient > 0.5 and

FDR < 0.05 were considered. GO term enrichment analysis of the putative target genes was subsequently performed using the *topGO* package(Alexa and Rahnenführer).

Building human protein–protein interactome and drug–target network

The human protein-protein interactome (PPI) and drug-target network were constructed in our previous studies[46–50]. In this study, the human PPI was constructed by integrating data from multiple established PPI databases supported by experimental evidence. This includes binary PPIs identified by high-throughput yeast-two-hybrid (Y2H) experiments(Luck et al. 2020) and kinase-substrate interactions derived from both low- and high-throughput literature sources, including the Human Protein Resource Database (HPRD)(Peri et al. 2004), dbPTM 3.0(Lu et al. 2013), Phospho.ELM(Dinkel et al. 2011), KinomeNetworkX(Cheng et al. 2014), PhosphoNetworks(Hu et al. 2014), and PhosphositePlus(Hornbeck et al. 2015). Additionally, signaling networks were curated from literature-based low-throughput experiments provided by SignaLink2.0(Fazekas et al. 2013). Protein complex interactions, comprising approximately 56,000 candidate interactions, were identified using robust affinity purification-mass spectrometry data from BioPlexV2.0(Huttlin et al. 2015). Furthermore, literature-curated PPIs derived from affinity purification coupled with mass spectrometry were included from sources such as HPRD(Goel et al. 2012), PINA(Cowley et al. 2012), MINT(Licata et al. 2012), InnateDB(Breuer et al. 2013), IntAct(Orchard et al. 2014), Instruct(Meyer et al. 2013), and BioGRID(Chatr-aryamontri et al. 2015). In total, this dataset encompasses 351,444 PPIs involving 17,706 protein nodes. The complete dataset is publicly available at https://alzgps.lerner.ccf.org.

The drug–target network was developed by integrating multiple reputable data sources, including DrugBank (version 4.3)(Law et al. 2014), BindingDB(Liu et al. 2007), ChEMBL (version 20)(Gaulton et al. 2012), the Therapeutic Target Database(Yang et al. 2016), PharmGKB(Whirl-Carrillo et al. 2012), and the IUPHAR/BPS Guide to PHARMACOLOGY(Pawson et al. 2014). Drug–target interactions were filtered using a cutoff of binding affinities (Ki, Kd, IC50, or EC50) ≤ 10 μM to ensure high confidence in the associations following the previous research(Zhou et al. 2023). The final network comprises a total of 29,934 interactions involving 7,407 drugs. The networks in this study were visualized with Cytoscape (version 3.10.2) (Shannon et al. 2003).

Network proximity-based drug repurposing

We first employed the closest network proximity method to identify potential drug repurposing candidates, as reported previously (Zhou et al. 2023; Xu et al. 2021). The closest distance between two sets of nodes in the protein–protein interaction network, such as the drug target set (X) and a disease-related gene set (Y), was defined as follows:

$$d_{(XY)} = \frac{1}{||X||+||Y||} \left(\sum_{x\in X} min_{y\in Y} d(x,y) + \sum_{y\in Y} min_{x\in X} d(x,y)\right) \quad (2)$$

where d(x,y) represents the shortest path length between protein x in X and protein y in Y within the human protein–protein interactome. To assess the significance of the observed proximity, we calculated a Z-score for $d_{(X,Y)}$ using a permutation-based approach. Random protein sets with similar degree distributions to X and Y were generated 1,000 times. For each random experiment, we computed the mean and standard deviation of the shortest distances. The Z-score was then calculated as:

$$Z_{d(XY)} = \frac{d(XY)-\mu_d}{\sigma_d} \quad (3)$$

Lastly, the results were filtered using thresholds of Z-score < −2 and FDR < 0.05 to identify significant drug candidates for further investigation.

Gene set enrichment analysis (GSEA)

We employed a GSEA-based approach for drug repurposing, as reported previously(Xu et al. 2021; Zhou et al. 2020). Initially, drug-gene signatures were obtained from the Connectivity Map (CMap) database(Lamb et al. 2006a; Subramanian et al. 2017; Lamb et al. 2006b), a comprehensive pharmacogenomic resource containing gene expression profiles from diverse human cell lines treated with thousands of small molecules, including both approved drugs and experimental compounds. Using the GSEA algorithm, we predicted enriched drugs for each cell type by leveraging the identified AD-associated differentially expressed genes as input. We then computed the enrichment score (ES) for each drug present in both the CMap database and the drug-target network using previously established methods(Xu et al. 2021; Zhou et al. 2020). The ES reflects the drug's potential to reverse the gene expression patterns within the given network(Xu et al. 2024a). To ensure statistical significance and meaningful positive enrichment, the results were filtered to retain only those with a FDR < 0.05 and an ES > 0. This enabled the prioritization of candidate compounds with potential therapeutic relevance for AD.

References

Alexa A, Rahnenführer J. Gene set enrichment analysis with topGO.

Bellenguez C, Küçükali F, Jansen IE, Kleineidam L, Moreno-Grau S, Amin N, Naj AC, Campos-Martin R, Grenier-Boley B, Andrade V, et al. 2022. New insights into the genetic etiology of Alzheimer's disease and related dementias. *Nat Genet* **54**: 412–436.

Breuer K, Foroushani AK, Laird MR, Chen C, Sribnaia A, Lo R, Winsor GL, Hancock REW, Brinkman FSL, Lynn DJ. 2013. InnateDB: systems biology of innate immunity and beyond—recent updates and continuing curation. *Nucleic Acids Research* **41**: D1228–D1233.

Chatr-aryamontri A, Breitkreutz B-J, Oughtred R, Boucher L, Heinicke S, Chen D, Stark C, Breitkreutz A, Kolas N, O'Donnell L, et al. 2015. The BioGRID interaction database: 2015 update. *Nucleic Acids Research* **43**: D470–D478.

Cheng F, Desai RJ, Handy DE, Wang R, Schneeweiss S, Barabási A-L, Loscalzo J. 2018. Network-based approach to prediction and population-based validation of in silico drug repurposing. *Nat Commun* **9**: 2691.

Cheng F, Jia P, Wang Q, Zhao Z. 2014. Quantitative network mapping of the human kinome interactome reveals new clues for rational kinase inhibitor discovery and individualized cancer therapy. *Oncotarget* **5**: 3697–3710.

Cheng F, Kovács IA, Barabási A-L. 2019a. Network-based prediction of drug combinations. *Nat Commun* **10**: 1197.

Cheng F, Lu W, Liu C, Fang J, Hou Y, Handy DE, Wang R, Zhao Y, Yang Y, Huang J, et al. 2019b. A genome-wide positioning systems network algorithm for in silico drug repurposing. *Nat Commun* **10**: 3476.

Cowley MJ, Pinese M, Kassahn KS, Waddell N, Pearson JV, Grimmond SM, Biankin AV, Hautaniemi S, Wu J. 2012. PINA v2.0: mining interactome modules. *Nucleic Acids Research* **40**: D862–D865.

Crowell HL, Soneson C, Germain P-L, Calini D, Collin L, Raposo C, Malhotra D, Robinson MD. 2020. muscat detects subpopulation-specific state transitions from multi-sample multi-condition single-cell transcriptomics data. *Nat Commun* **11**: 6077.

Dinkel H, Chica C, Via A, Gould CM, Jensen LJ, Gibson TJ, Diella F. 2011. Phospho.ELM: a database of phosphorylation sites—update 2011. *Nucleic Acids Research* **39**: D261–D267.

Emani PS, Liu JJ, Clarke D, Jensen M, Warrell J, Gupta C, Meng R, Lee CY, Xu S, Dursun C, et al. 2024. Single-cell genomics and regulatory networks for 388 human brains. *Science* **384**: eadi5199.

Fazekas D, Koltai M, Türei D, Módos D, Pálfy M, Dúl Z, Zsákai L, Szalay-Bekő M, Lenti K, Farkas IJ, et al. 2013. SignaLink 2 – a signaling pathway resource with multi-layered regulatory networks. *BMC Systems Biology* **7**: 7.

Gabitto MI, Travaglini KJ, Rachleff VM, Kaplan ES, Long B, Ariza J, Ding Y, Mahoney JT, Dee N, Goldy J, et al. 2024. Integrated multimodal cell atlas of Alzheimer's disease. *Nat Neurosci* 1–18.

Gaulton A, Bellis LJ, Bento AP, Chambers J, Davies M, Hersey A, Light Y, McGlinchey S, Michalovich D, Al-Lazikani B, et al. 2012. ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Research* **40**: D1100–D1107.

Goel R, Harsha HC, Pandey A, Prasad TSK. 2012. Human Protein Reference Database and Human Proteinpedia as resources for phosphoproteome analysis. *Mol BioSyst* **8**: 453–463.

Granja JM, Corces MR, Pierce SE, Bagdatli ST, Choudhry H, Chang HY, Greenleaf WJ. 2021. ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis. *Nat Genet* **53**: 403–411.

Heumos L, Schaar AC, Lance C, Litinetskaya A, Drost F, Zappia L, Lücken MD, Strobl DC, Henao J, Curion F, et al. 2023. Best practices for single-cell analysis across modalities. *Nat Rev Genet* 1–23.

Hornbeck PV, Zhang B, Murray B, Kornhauser JM, Latham V, Skrzypek E. 2015. PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Research* **43**: D512–D520.

Hu J, Rho H-S, Newman RH, Zhang J, Zhu H, Qian J. 2014. PhosphoNetworks: a database for human phosphorylation networks. *Bioinformatics* **30**: 141–142.

Huttlin EL, Ting L, Bruckner RJ, Gebreab F, Gygi MP, Szpyt J, Tam S, Zarraga G, Colby G, Baltier K, et al. 2015. The BioPlex Network: A Systematic Exploration of the Human Interactome. *Cell* **162**: 425–440.

Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ, Lerner J, Brunet J-P, Subramanian A, Ross KN, et al. 2006a. The Connectivity Map: Using Gene-Expression Signatures to Connect Small Molecules, Genes, and Disease. *Science* **313**: 1929–1935.

Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ, Lerner J, Brunet J-P, Subramanian A, Ross KN, et al. 2006b. The Connectivity Map: Using Gene-Expression Signatures to Connect Small Molecules, Genes, and Disease. *Science* **313**: 1929–1935.

Law V, Knox C, Djoumbou Y, Jewison T, Guo AC, Liu Y, Maciejewski A, Arndt D, Wilson M, Neveu V, et al. 2014. DrugBank 4.0: shedding new light on drug metabolism. *Nucleic Acids Research* **42**: D1091–D1097.

Licata L, Briganti L, Peluso D, Perfetto L, Iannuccelli M, Galeota E, Sacco F, Palma A, Nardozza AP, Santonico E, et al. 2012. MINT, the molecular interaction database: 2012 update. *Nucleic Acids Research* **40**: D857–D861.

Liu T, Lin Y, Wen X, Jorissen RN, Gilson MK. 2007. BindingDB: a web-accessible database of experimentally determined protein–ligand binding affinities. *Nucleic Acids Research* **35**: D198–D201.

Lu C-T, Huang K-Y, Su M-G, Lee T-Y, Bretaña NA, Chang W-C, Chen Y-J, Chen Y-J, Huang H-D. 2013. dbPTM 3.0: an informative resource for investigating substrate site specificity and functional association of protein post-translational modifications. *Nucleic Acids Research* **41**: D295–D305.

Luck K, Kim D-K, Lambourne L, Spirohn K, Begg BE, Bian W, Brignall R, Cafarelli T, Campos-Laborie FJ, Charloteaux B, et al. 2020. A reference map of the human binary protein interactome. *Nature* **580**: 402–408.

Mathys H, Peng Z, Boix CA, Victor MB, Leary N, Babu S, Abdelhady G, Jiang X, Ng AP, Ghafari K, et al. 2023. Single-cell atlas reveals correlates of high cognitive function, dementia, and resilience to Alzheimer's disease pathology. *Cell* **186**: 4365-4385.e27.

McInnes L, Healy J, Melville J. 2020. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. http://arxiv.org/abs/1802.03426 (Accessed November 20, 2024).

Meyer MJ, Das J, Wang X, Yu H. 2013. INstruct: a database of high-quality 3D structurally resolved protein interactome networks. *Bioinformatics* **29**: 1577–1579.

Morabito S, Miyoshi E, Michael N, Shahin S, Martini AC, Head E, Silva J, Leavy K, Perez-Rosendahl M, Swarup V. 2021. Single-nucleus chromatin accessibility and transcriptomic characterization of Alzheimer's disease. *Nat Genet* **53**: 1143–1155.

Ober-Reynolds B, Wang C, Ko JM, Rios EJ, Aasi SZ, Davis MM, Oro AE, Greenleaf WJ. 2023. Integrated single-cell chromatin and transcriptomic analyses of human scalp identify gene-regulatory programs and critical cell types for hair and skin diseases. *Nat Genet* **55**: 1288–1300.

Orchard S, Ammari M, Aranda B, Breuza L, Briganti L, Broackes-Carter F, Campbell NH, Chavali G, Chen C, del-Toro N, et al. 2014. The MIntAct project—IntAct as a common curation platform for 11 molecular interaction databases. *Nucleic Acids Research* **42**: D358–D363.

Pawson AJ, Sharman JL, Benson HE, Faccenda E, Alexander SPH, Buneman OP, Davenport AP, McGrath JC, Peters JA, Southan C, et al. 2014. The IUPHAR/BPS Guide to PHARMACOLOGY: an expert-driven knowledgebase of drug targets and their ligands. *Nucleic Acids Res* **42**: D1098-1106.

Peri S, Navarro JD, Kristiansen TZ, Amanchy R, Surendranath V, Muthusamy B, Gandhi TKB, Chandrika KN, Deshpande N, Suresh S, et al. 2004. Human protein reference database as a discovery resource for proteomics. *Nucleic Acids Research* **32**: D497–D501.

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. 2003. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res* **13**: 2498–2504.

Subramanian A, Narayan R, Corsello SM, Peck DD, Natoli TE, Lu X, Gould J, Davis JF, Tubelli AA, Asiedu JK, et al. 2017. A Next Generation Connectivity Map: L1000 platform and the first 1,000,000 profiles. *Cell* **171**: 1437-1452.e17.

Traag VA, Waltman L, van Eck NJ. 2019. From Louvain to Leiden: guaranteeing well-connected communities. *Sci Rep* **9**: 5233.

van Dijk D, Sharma R, Nainys J, Yim K, Kathail P, Carr AJ, Burdziak C, Moon KR, Chaffer CL, Pattabiraman D, et al. 2018. Recovering Gene Interactions from Single-Cell Data Using Data Diffusion. *Cell* **174**: 716-729.e27.

Weirauch MT, Yang A, Albu M, Cote AG, Montenegro-Montero A, Drewe P, Najafabadi HS, Lambert SA, Mann I, Cook K, et al. 2014. Determination and inference of eukaryotic transcription factor sequence specificity. *Cell* **158**: 1431–1443.

Whirl-Carrillo M, McDonagh EM, Hebert JM, Gong L, Sangkuhl K, Thorn CF, Altman RB, Klein TE. 2012. Pharmacogenomics Knowledge for Personalized Medicine. *Clinical Pharmacology & Therapeutics* **92**: 414–417.

Wolf FA, Angerer P, Theis FJ. 2018. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biology* **19**: 15.

Wolock SL, Lopez R, Klein AM. 2019. Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. *Cell Syst* **8**: 281-291.e9.

Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, Feng T, Zhou L, Tang W, Zhan L, et al. 2021. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Camb)* **2**: 100141.

Xiong X, James BT, Boix CA, Park YP, Galani K, Victor MB, Sun N, Hou L, Ho L-L, Mantero J, et al. 2023. Epigenomic dissection of Alzheimer's disease pinpoints causal variants and reveals epigenome erosion. *Cell* **186**: 4422-4437.e21.

Xu J, Mao C, Hou Y, Luo Y, Binder JL, Zhou Y, Bekris LM, Shin J, Hu M, Wang F, et al. 2022. Interpretable deep learning translation of GWAS and multi-omics findings to identify pathobiology and drug repurposing in Alzheimer's disease. *Cell Reports* **41**: 111717.

Xu J, Song W, Xu Z, Danziger MM, Karavani E, Zang C, Chen X, Li Y, Paz IMR, Gohel D, et al. 2024a. Single-microglia transcriptomic transition network-based prediction and real-world patient data validation identifies ketorolac as a repurposable drug for Alzheimer's disease. *Alzheimers Dement*.

Xu J, Zhang P, Huang Y, Zhou Y, Hou Y, Bekris LM, Lathia J, Chiang C-W, Li L, Pieper AA, et al. 2021. Multimodal single-cell/nucleus RNA sequencing data analysis uncovers molecular networks between disease-associated microglia and astrocytes with implications for drug repurposing in Alzheimer's disease. *Genome Res* **31**: 1900–1912.

Xu S, Hu E, Cai Y, Xie Z, Luo X, Zhan L, Tang W, Wang Q, Liu B, Wang R, et al. 2024b. Using clusterProfiler to characterize multiomics data. *Nat Protoc* **19**: 3292–3320.

Yang H, Qin C, Li YH, Tao L, Zhou J, Yu CY, Xu F, Chen Z, Zhu F, Chen YZ. 2016. Therapeutic target database update 2016: enriched resource for bench to clinical drug target and targeted pathway information. *Nucleic Acids Research* **44**: D1069–D1074.

Yu G, Wang L-G, Han Y, He Q-Y. 2012. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS* **16**: 284–287.

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, et al. 2008. Model-based Analysis of ChIP-Seq (MACS). *Genome Biology* **9**: R137.

Zhou Y, Hou Y, Shen J, Huang Y, Martin W, Cheng F. 2020. Network-based drug repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2. *Cell Discov* **6**: 14.

Zhou Y, Liu Y, Gupta S, Paramo MI, Hou Y, Mao C, Luo Y, Judd J, Wierbowski S, Bertolotti M, et al. 2023. A comprehensive SARS-CoV-2–human protein–protein interactome reveals COVID-19 pathobiology and potential host therapeutic targets. *Nat Biotechnol* **41**: 128–139.

126. Perez-Gonzalez, A.P., et al., The ROSMAP project: aging and neurodegenerative diseases through omic sciences. *Front Neuroinform*, 2024. **18**: 1443865.