S.T. Yau High School Science Award

Research Report

The Team

Name of team member: Cylina Yuechen Wang
School: Loomis Chaffee School
City, Country: Connecticut, USA

Name of supervising teacher: Zhaoyang Huang
Job Title: Director of the Electrophysiology Department, Department of Neurology, Xuanwu Hospital, Associate Chief Physician
School/Institution: Xuanwu Hospital of Capital Medical University
City, Country: Beijing, China

Title of Research Report

HTT-HDAC Axis links Major Depressive Disorder and Alzheimer's Disease: A Transformer-Based Multi-Omics Dissection

Date

August 19, 2025

HTT-HDAC Axis links Major Depressive Disorder and Alzheimer's Disease: A Transformer-Based Multi-Omics Dissection

Cylina Yuechen Wang

Abstract

Alzheimer's disease (AD), affecting over 55 million globally, is marked by cognitive decline, amyloid-β plaques, and tau tangles. Major depressive disorder (MDD), impacting 280 million worldwide, involves persistent depression and anhedonia. Strikingly, 30–50% of AD patients develop MDD, while late-life MDD increases AD risk by 1.5–2 fold, suggesting a bidirectional relationship between the two. Despite evidence of shared mechanisms, the links between AD and MDD—spanning genetic, molecular, and clinical dimensions—remain poorly defined, largely due to conflicting genetic evidence and the inconsistent efficacy of antidepressants in treating comorbid populations. Here, we identified a gene regulatory network (GRN) linking AD and MDD, established a transformer model to identify convergent pathways underlying their comorbidity, and evaluated the established network, prioritized genes, and drug candidates in *C. elegans* transgenic models. Contrary to prior studies that focused on genetic or pharmacological mechanisms independently, this study explores the comorbidity between AD and MDD through multiple dimensions. Our findings reveal that this comorbidity is driven by shared biological pathways, particularly chromatin organization and epigenetic regulation, centered on the GRN we identified. These findings challenge the traditional view of AD and MDD comorbidity as a shared pathophysiological process instead of a complex network. By linking computational predictions with experimental validation, this study provides a mechanistic framework for future dual-target therapies. This advances the precision of treatment strategies for comorbid AD and MDD, addressing a critical gap in neurodegenerative and psychiatric knowledge and care.

Keywords: Alzheimer's disease, major depressive disorder, single-cell transcriptomics, transformer model, machine learning, large language model, epigenetics

Acknowledgements

This work was inspired, in part, by personal encounters that deepened my understanding of aging and its emotional and neurological sophistication. I still recall seeing Mr. Li, a former neighbor of ours, shredding a piece of paper with shaking hands. I saw how a man, once calm and warm, now muttered curses at sparrows. When I later learned that he had been diagnosed with Alzheimer's, my grandmother added, "But he's been different… long before the memory loss." Years later, as I ponder and delve into scRNA-seq data and uncover links between neurodegenerative and neuropsychiatric disorders, I began to understand how depression and personality changes might be rooted in early molecular shifts in neurodegeneration.

I am very grateful for my mentor, Dr. Zhaoyang Huang at Xuanwu Hospital. Under his guidance, I explored the connection between AD and MDD. The project evolved from polygenic risk scoring to gene prioritization, and finally to network-based modeling. I learned how to process and analyze GWAS, scRNA-seq, and ATAC-seq data, and gained critical knowledge about LLMs and transformer models from Dr. Huang. Dr. Huang also guided me through practical techniques, including plasmid design, dsRNA binding validation, *C. elegans* synchronization, and survival analyses. His ongoing work linking EEG signals to pathology classification helped shape my approach to modeling transcriptomic data.

To Mom and Dad, who video-called every Saturday night while I squinted at Visual Studio Code and passionately explained the concepts of non-linear dependencies and gene-pathway matrices. You listened as I ranted about "attention mechanisms lay the foundation of transformer models". Through your patience and willingness to listen, you reminded me of a world beyond the lab.

To my MacBook, who withstood the 2AM all-nighters, the spilled coffee, the 76 Chrome tabs of NCBI and GitHub pages. When the PFC MDD data flooded your 256GB SSD, I was determined to make space for yet another expression matrix by deleting WeChat and my photos.

Lastly, I would like to thank Mr. George Osei-Mensah (The Loomis Chaffee School) for supporting this project, Mr. Zhifei Li (Tsinghua International School Daoxiang-Lake) for laying the foundation of my biology knowledge through his engaging teaching, and my family for their unwavering support.

Academic Integrity Statement

I hereby declare that the paper submitted is comprised of original research and results obtained under the guidance of my mentor, Dr. Zhaoyang Huang. All single-cell transcriptomic analyses and pathway analyses data are retrieved from publicly available databases: NCBI GEO Accession, Allen Brain Atlas, SEA-AD, and the Psychiatric Genomics Consortium (PGCO). All computational analyses are conducted by the contestant using RStudio, Visual Studio Code, UCSC Genome Browser, QIAGEN Ingenuity Pathway Analysis (IPA), and STRING. Drug and mechanism validation using *C. elegans* model are conducted under Dr. Huang's guidance. It is important to note that there is no direct collaboration or professional relationship between Dr. Huang and the Loomis Chaffee School. To the best of my knowledge, the paper does not include research results that have been published or written by others, except for those specifically marked and acknowledged in the article. If there are any inaccuracies, I am willing to assume all relevant responsibilities.

Signature of the Contestant: Cylina Wang

Date: 2025/06/27

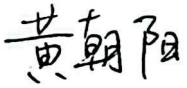
Declaration of Academic Integrity

The participating team declares that the paper submitted is comprised of original research and results obtained under the guidance of the instructor. To the team's best knowledge, the paper does not contain research results, published or not, from a person who is not a team member, except for the content listed in the references and the acknowledgment. If there is any misinformation, we are willing to take all the related responsibilities.

Name of team member: Yuechen Wang

Signatures of team member: 

Name of the instructor: Zhaoyang Huang

Signature of the instructor: 

Date: 2025/7/13

Table of Contents

1. Introduction **1**
1.1 Epidemiological Link Between AD and MDD 1
1.2 Pathophysiological Overlap between AD and MDD 2
1.3 Pharmacological and Clinical Connection between AD and MDD 2
1.4 Applying Transformer Architectures to Map Links between AD and MDD 3
1.5 Innovation and Significance 3
1.6 Research Question and Hypothesis 4
2. Materials and Methods **5**
2.1 Multi-Omic Data Acquisition and Preprocessing 5
2.2 Construction of QuokkaVision 6
2.3 Validation of QuokkaVision 10
2.4 Functional and In Vivo Validation of Candidate Genes in C. elegans 11
3. Results **14**
3.1 Integrated PPI Network Reveals a Htt-HDAC Axis in MDD and AD 14
3.2 Htt as a Transcriptional Hub Linking Epigenetics and Synaptic Dysfunction 15
3.3 Transcriptomic Profiling Highlights Neurodevelopmental, GABAergic, Histone Dysregulation Across Subtypes 16
3.4 Chromatin Accessibility Shifts at the Htt Locus in AD Neurons 18
3.5 QuokkaVision Links Htt-HDAC Network to Layer-Specific Dysregulation 19
3.6 Functional Validation of the Htt-HDAC Network via dsRNAi Suppression 21
3.7 Therapeutic Efficacy Evaluation via Pathological Inhibition of the Htt-HDAC Network by HDAC Inhibitors 22
4. Discussion **24**
4.1 QuokkaVision in Context: From Architecture to Implications 24
4.2 Implications of the Htt-HDAC Axis 25
4.3 Limitations and Future Directions 26
5. Conclusion and Perspectives **27**
Conflicts of Interest **27**
Data Availability Statement **28**
Supplemental Data **29**
References **36**

1. Introduction

Age-related cognitive impairment often comes with emotional alterations such as sadness, depression, and anxiety. However, the biological mechanism behind this remains poorly understood, limiting the development of effective interventions in elder care. Alzheimer's disease (AD), the most prevalent form of dementia, is marked not only by cognitive decline but also by emotional changes such as irritability and persistent sadness (Heilman & Nadeau, 2022). These symptoms often emerge before or alongside cognitive deficits, which suggests a more complicated overlap (Heilman & Nadeau, 2022). In recent years, major depressive disorder (MDD) has been increasingly diagnosed with AD. MDD is a neuropsychiatric condition characterized by persistent sadness and altered behavioral regulation (Cui et al., 2024). This overlap raises a critical question: **Is there a shared neurodegenerative mechanism between AD and MDD on a genetic and molecular level?** To address this question, we hypothesize that there exists a gene regulatory network between AD and MDD. Below, three perspectives of the link between AD and MDD will be examined: epidemiological link, pathophysiological overlap, and pharmacological connection.

1.1 Epidemiological Link Between AD and MDD

Depression is both a risk factor and a common neuropsychiatric symptom observed in AD. Within the 55 million AD patient population, up to 60% exhibit at least one neuropsychiatric symptom (Lyketsos et al., 2002), and approximately 38-40% of AD patients experience comorbid depression (Botto et al., 2022). This is a prevalence far exceeding chance association.

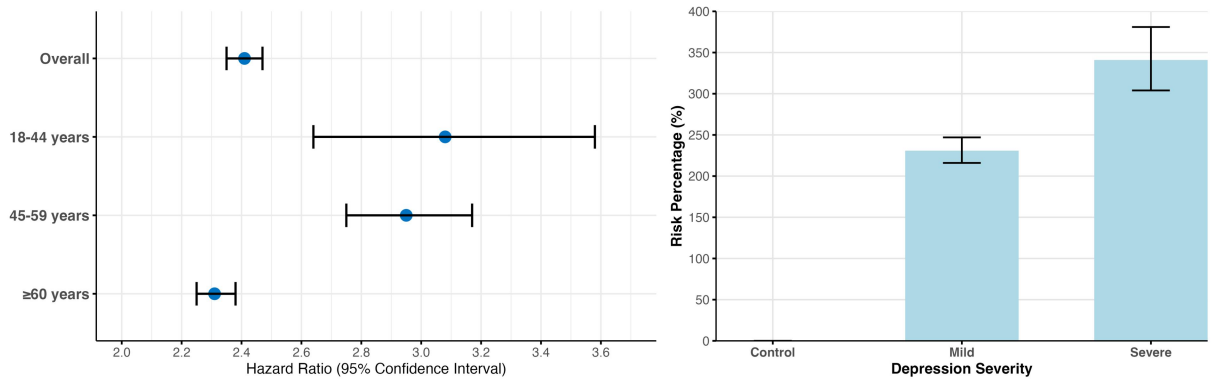


Figure 1. Risk of AD in Depressed Populations. (a) Hazard Ratio of AD Among Individuals with Depression, stratified by age ranges; Data adapted from Heun et al. (2001). (b) Based on Depression Severity; Data adapted from Kim et al. (2021).

This association is also supported by a severity-dependent observation: longitudinal data show that depression severity correlates with dementia risk (**Figure 1b**) (Holmquist et al., 2020), and AD-related neurodegeneration may itself trigger depressive symptoms. For example, 16-32% of AD patients experience depressive symptoms, which, at any stage, accelerates disease progression (Aguera-Ortiz et al., 2021; Asmer et al., 2018). It is also shown that late life MDD

patients suffer an approximately doubled risk of AD development (**Figure 1a**) (Barnes et al., 2012). These findings strongly suggest a shared mechanism between AD and MDD, with MDD as both a risk factor for AD and a secondary consequence of AD.

1.2 Pathophysiological Overlap between AD and MDD

In AD, brain regions, such as the hippocampus, amygdala, entorhinal cortex, and temporal lobe, exhibit significant atrophy and neuronal loss. Interestingly, these areas are also involved in emotional regulation and processing. This anatomical overlap suggests the possibility of neurodegeneration linked to emotional dysregulation and partially answers how AD and MDD are pathophysiologically connected. Evidence increasingly supports that depression may be both a risk factor and an early symptom of AD.

The directionality of this relationship remains less certain. For one reason, the heterogeneity of depressive phenotypes complicates the mechanistic study of their connection (Monereo-Sanchez et al., 2021). Clinical and neuroimaging data were able to suggest shared pathology but remained inconclusive regarding the possible mechanisms linking AD and MDD. For instance, both diseases often exhibit common characteristics such as the reduction of hippocampal volume and amyloid-β (Aβ) accumulation, which may together exacerbate depressive symptoms (Babulal et al., 2020; Pomara et al., 2012).

Similarly, genetic studies remain inconclusive. While one study found no common polygenic structure between AD and MDD (Gibson et al., 2017), others have identified approximately 98 causal genetic variants that overlaps between AD and MDD and linked higher AD polygenic risk scores with increased risk of late-life depression (Monereo-Sanchez et al., 2021; Wingo et al., 2023). These contradictories warrant further investigation. This study aims to resolve these inconsistencies to pave way for targeted therapeutic interventions that address the shared and distinct pathological pathways of both disorders.

1.3 Pharmacological and Clinical Connection between AD and MDD

Pharmacological evidence further strengthens the connection between AD and MDD. Although primarily prescribed for depression, selective serotonin reuptake inhibitors (SSRIs) exhibit unexpected benefits in AD models. Prolonged SSRI use of approximately three years delays AD progression in individuals with mild cognitive impairment (MRI) (Bartels et al., 2018). Studies also find that escitalopram, a type of SSRI, reduces Aβ plaques in cognitively normal older adults (Cirrito et al., 2020; Sheline et al., 2020). These findings suggest serotonin signaling and other potential mechanisms linking MDD with AD.

However, not all antidepressants yield positive effects. Other classes of antidepressants, including norepinephrine reuptake inhibitors (SNRIs) and tricyclic antidepressants (TCAs), were found to increase the risk of dementia among adults of middle or older age (Brauer et al., 2019). This inconsistency in antidepressant efficacy suggests the central role of target

specificity to pathways linking AD and MDD in determining the therapeutic efficacy of the treatment (Dudas et al., 2018; Orgeta et al., 2017).

Ongoing clinical trials further support the potential of shared pathways between AD and MDD. One study, *Depression Treatment and Aβ Dynamics*, conducted at the New York University School of Medicine, is evaluating the impact of escitalopram oxalate on Alzheimer's biomarkers in cognitive unimpaired older adults with MDD (*Aβ Dynamics in LLMD*, 2021). Along this line, a Transcranial Magnetic Stimulation for MCI trial, sponsored by the National Institute on Aging (NIA), is testing intermittent theta-burst repetitive transcranial magnetic stimulation as a non-pharmacological treatment for individuals with both MCI and depression (*Transcranial Magnetic Stimulation for MCI (PUSH2)*, 2023). These pharmacological and neuromodulatory findings support a mechanism linking AD and MDD and the therapeutic potential of interventions that target both disorders.

1.4 Applying Transformer Architectures to Map Links between AD and MDD

A persistent challenge in neuropsychiatric research remains in resolving the multi-layer genetic, molecular, and cellular interactions. Differential expression approaches often have limited capability due to normalization methods and cumulative biases. Essentially, these methods force non-linear gene interactions into linear frameworks. These limitations are especially pronounced when analyzing single-cell genomics data with varied cellular heterogeneity, which is the case for neurons in the brain, where preserving biological integrity and scaling to millions of data points are equally important.

To address this challenge, researchers have turned to artificial intelligence (AI), particularly the transformer model, which is a neural network architecture originally developed for processing sequential data. Unlike traditional models that analyze data step by step, transformers use self-attention to weigh the importance of each element in a sequence relative to others, allowing them to capture long-range and context-dependent relationships. This mechanism not only allows efficient parallel processing but also enables large language models (LLMs), like ChatGPT, in prioritizing subtle signals across datasets. In recent years, transformer models have been increasingly used for researchers to identify unconventional relationships across genes, pathways, and cell types (Vaswani et al., 2017). In biomedical contexts, transformers have demonstrated success in predicting drug-target interactions (Huang et al., 2021), decoding genomic sequence (Dalla-Torre et al., 2025), and detecting disease patterns (Vorontsov et al., 2024). Building on these advances, our study aims to use the ability of transformer models to detect long-range dependencies to reveal previously overlooked pathways linking AD and MDD.

1.5 Innovation and Significance

To the best of our knowledge, this study is one of the first to integrate GWAS, single-cell transcriptomics, and chromatin accessibility data to map shared molecular pathways between AD and MDD. We also developed QuokkaVision, a transformer-based framework that

resolves nonlinear gene-pathway-subtype interactions, such as *Htt*'s chromatin scaffolding role in hippocampal neurons.

In 2019, the World Alzheimer's Report revealed that more than 55 million individuals globally were affected by dementia, with the associated economic burden estimated at 1.3 trillion (*World Alzheimer Report 2024: Global changes in attitudes to dementia*, 2024). This disease burden is particularly evident in comorbid disorders, particularly AD and MDD. Individually, AD is ranked seventh-leading cause of death, and MDD as third cause of disease burden (Bains & Abdijadid, 2025). However, this report doesn't take comorbid patients into account, who suffer significant disease burden.

Our identification of shared pathways lays the foundation of dual-target interventions that simultaneously address neurodegenerative and neuropsychiatric conditions. For example, one promising candidate is histone deacetylases (HDACs) inhibitors. HDACs are enzymes that regulate gene expression epigenetically by condensing chromatin, which may suppress genes critical for plasticity and neuronal survival (Milazzo et al., 2020). In contrast, HDAC inhibitors block this process, and reactivate suppressed gene networks (Milazzo et al., 2020). Compounds like suberoylanilide hydroxamic acid (SAHA; otherwise known as vorinostat) have been shown preclinically to reduce amyloid toxicity (AD) and enhance plasticity (MDD & AD) (Kilgore et al., 2010; Meng et al., 2014). Together, these strategies together offer a hopeful path to targeting AD and MDD comorbidity in aging populations.

1.6 Research Question and Hypothesis

This study aims to explore the underlying genetic and molecular pathways between AD and MDD through an integrative approach which adopts an MDD-to-AD framework. This framework allows us to explore a more clearly defined disease progression as the inherent heterogeneity of MDD complicates the process of associating AD with MDD.

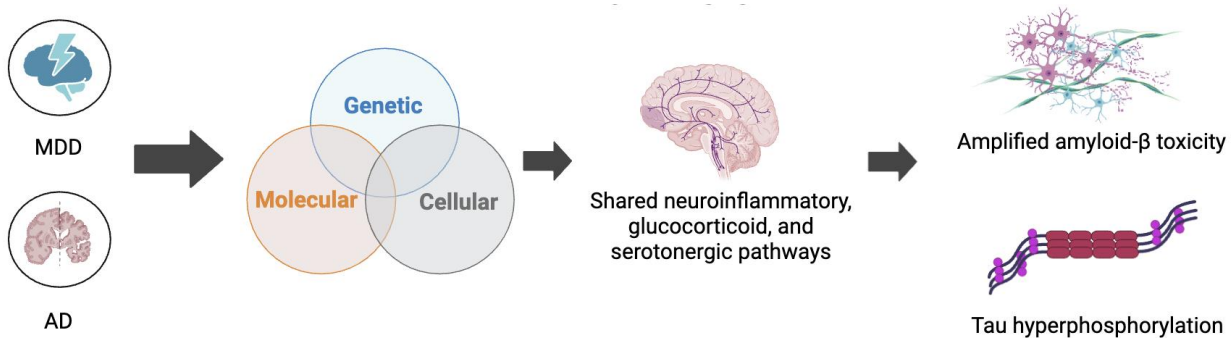


Figure 2. Hypothesized molecular framework linking MDD with AD (created with Biorender).

We hypothesize that there exists a gene regulatory network between AD and MDD that would influence certain pathways, such as neuroinflammatory, glucocorticoid, and serotonergic pathways, and amplify Aβ toxicity and tau hyperphosphorylation (**Figure 2**).

2. Materials and Methods

In this study, multi-omic datasets were processed and analyzed to reveal the underlying link between AD and MDD (**Figure 3**). To reveal shared protein networks, overlapping genes between AD and MDD were identified from open-sourced genome-wide association study (GWAS) and single-cell RNA sequencing (scRNA-seq) datasets and subsequently analyzed via STRING. To trace underlying pathways, a transformer model, namely QuokkaVision, was developed to identify the critical yet unconventional pathways linking the critical gene(s). Based on these findings, targeted pharmacological treatments were selected and functionally tested in AD *Caenorhabditis elegans* (*C. elegans*) models carrying RNAi knockouts of prioritized gene(s). Through in vivo validation, the validity of the link and efficacy of the treatment will be assessed through a multitude of assays. These experiments aim to identify potential dual-target therapeutics for both AD and MDD.

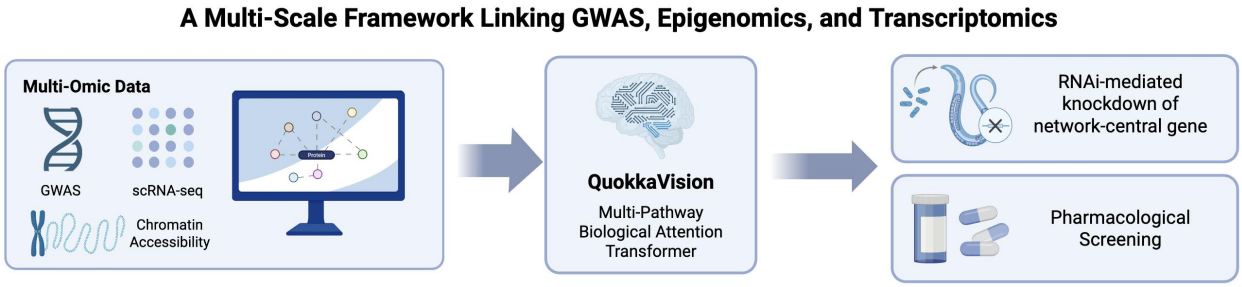


Figure 3. Integrated multi-omic workflow for discovery and validation (created with Biorender).

2.1 Multi-Omic Data Acquisition and Preprocessing

GWAS Datasets

The genetic variants associated with MDD were identified through a meta-analysis of large-scale GWAS consortia. To identify overlapping genes with genome-wide significance between MDD GWAS datasets, the PMC3837431 dataset of MDD mega-analysis from Psychiatric Genomics Consortium (n = 108,287 cases/controls) and the PMC6522363 dataset (n = 807,553 cases/controls), a genome-wide meta-analysis highlighting prefrontal cortex-linked loci and genes, were employed (Howard et al., 2019; Major Depressive Disorder Working Group of the Psychiatric et al., 2013). Protein-protein interaction (PPI) networks and KEGG/GO pathway enrichment analyses were then conducted via the STRING database (v12.0) to map shared biological pathways (Szklarczyk et al., 2023).

Single-cell RNA Datasets

scRNA-seq datasets were selected and obtained from publicly available repositories to capture neuronal heterogeneity in disease-specific contexts through statistical computation. The GSE208438 dataset, comprising scRNA-seq datasets of induced pluripotent stem cell (iPSC)-derived neurons from MDD and healthy control cohorts, was selected to assess pan-cellular

and cell-type-specific transcriptional shifts across cohorts in MDD (Lu et al., 2023). The GSE157827 dataset, comprising single-nucleus RNA sequencing of 169,496 nuclei from the prefrontal cortex (PFC) of AD patients and healthy controls, was also selected to conduct transcriptomic analyses on AD patients (Lau et al., 2020). Datasets underwent a multi-step preprocessing workflow to mitigate technical noise and batch effects. Cellular transcriptomes were filtered to exclude poor-quality cells. Normalization was performed using the log normalization method to standardize data representation, followed by data integration via principal component analysis (PCA) to harmonize batches. Dimensionality reduction was achieved through uniform manifold approximation and projection (UMAP) using the top 30 principal components (PCs). Ingenuity Pathway Analysis was applied to analyze GSE208438, centering on a regulatory network of prioritized gene(s).

Epigenetic Data

Chromatin accessibility at the *Htt* locus (chr4: 3,074,147–3,243,896, GRCh38) was analyzed using single-nucleus ATAC-seq (snATAC-seq) data from the Seattle Alzheimer's Disease Brain Cell Atlas (SEA-AD), under the Allen Brain Atlas. Preprocessed bigWig files representing normalized accessibility scores for individual neuronal subtypes in the PFC were directly retrieved from the SEA-AD portal. Cell-type-specific tracks included SST inhibitory neurons, L5 excitatory neurons, and L2/3 excitatory neurons. Accessibility scores for each neuronal subtype were visualized on the UCSC Genome Browser (GRCh38/hg38 assembly) using unmodified SEA-AD tracks. Aggregate accessibility signals at the *Htt* locus were computed by summing per-base accessibility scores across the locus for each cell type. Differential accessibility between AD (NC3 pathology) and neurotypical controls (NC0) was calculated.

Clustering Analysis and Cell Type Annotation Using Reference Gene Markers

Clustering was performed via an unsupervised machine learning algorithm on the shared nearest neighbor graph derived from harmonized PCs. Differentially expressed genes (DEGs) across clusters were identified using a likelihood-ratio test ($|\log2 \text{ fold-change}| > 0.5$, adjusted $p < 0.05$), and are visualized through volcano plots. Cell type identities were annotated using the following canonical marker genes: serotonergic neurons (*SLC6A4*, *TPH2*, *FEV*), astrocytes (*AQP4*, *GLUL*), glutamatergic neurons (*SLC17A6*, *SLC17A7*), oligodendrocyte precursor cells (*PDGFRA*, *VCAN*), excitatory neurons (*SLC17A7*, *CAMK2A*), and inhibitory neurons (*GAD1*, *GAD2*, *SLC32A1*).

2.2 Construction of QuokkaVision

QuokkaVision Model Architecture

The QuokkaVision model processes single-cell RNA-seq data through a pathway-informed transformer architecture (**Figure 4**). By integrating pathway activation scores with gene expression, this study demonstrates QuokkaVision's capacity to: (1) label cell types with

higher resolution than existing methods, (2) depict relevant genetic and regulatory pathways upregulated or downregulated across cell types via quantifiable pathway-gene covariance matrices, and (3) map disease-relevant genes to their respective genetic and biological pathways through entropy-optimized information transfer.

Structural-wise, QuokkaVision differentiates itself from prior models for several important reasons: (1) It's scalable. QuokkaVision adopts a memory-efficient structure that enables large-scale multi-disease analysis. (2) It's flexible. QuokkaVision enables the use of different pathway databases and demonstrates a wide variety of applications in terms of pathway analysis on the genetic, molecular, cellular, and human diseases levels. (3) It's unbiased. QuokkaVision processes cell-type classification independent of pathway analysis with independently trained gene-centric classifier and context-specific pathway activation patterns derived from self-attention mechanisms.

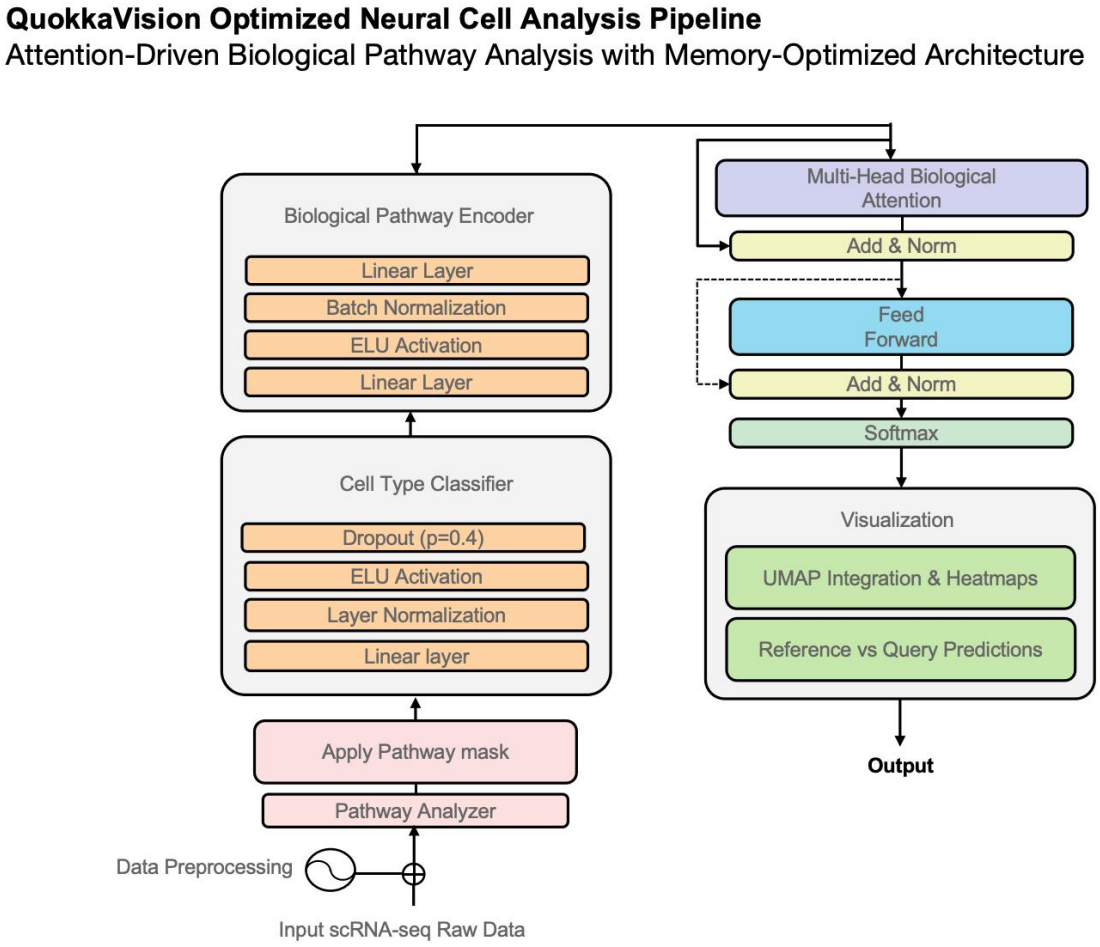


Figure 4. Algorithmic framework of QuokkaVision. QuokkaVision maximizes information extraction through multi-scale attention mechanisms.

Pathway Mask Construction

The pathway-gene Matrix M is constructed using the Allen Brain Atlas (Siletti et al., 2023) by mapping all genes in the input data to pathways from the Allen Brain Atlas and transposing and casting it to a binary tensor.

Cell-type Classifier

For each cell, normalized gene expression values (n genes) are first projected into a pathway-encoded embedding space. This is done using a knowledge-guided mask matrix M, which is a binary matrix of size (n × p), where p is the number of pathways. The transformation is applied as:

$$X_{pathways} = X \cdot M$$

where M is a binary matrix where each entry $M_{i,j} = 1$ if gene i is associated with pathway j, based on the data from the Allen Brain Atlas.

Next, the pathway embedding undergo processing in a biological encoder, using an exponential linear unit (ELU) activation function with layer normalization:

$$H = ELU(LayerNorm(X_{pathways}W_{enc} + b_{enc}))$$

where W_{enc} is a learnable weight matrix (p × 512) using backpropagation that updates the weight matrix to minimize the loss function using optimization algorithms, and b_{enc} is a learnable bias term.

The model then applies multi-head biological attention to compute contextualized pathway representations:

$$Q = K = V = H$$

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

$$H_{atten} = MultiHead(H) = Concatenate(head_1, \ldots, head_8)W_0$$

where Q, K, V are the query, key, and value matrices respectively, and d_k is the dimension per attention head (set to 64), and the attention mechanism utilizes 8 parallel attention heads.

Finally, the model classifies cell types using a fully connected layer with batch normalization and ELU activation:

$$Y = softmax(\quad FC_{256\to nc}(ELU(FC_{512\to 256}(H + H_{attn}))))$$

where FC denotes fully connected layers with batch normalization and nc is the number of cell types.

Data Processing Pipeline

1. Data Merging: The reference and query datasets are integrated by finding their common genes:

$$G_{common} = G_{ref} \cap G_{query}$$

2. Quality Control: Cells expressing fewer than 200 genes are removed; genes detected in fewer than 30 cells are filtered out; cells with mitochondrial gene content exceeding 20% are removed.

3. Normalization: Expression values are log-transformed.

Reference-Specific Model Training

The model trains exclusively on reference data through cross-entropy loss L2 regularization:

$$\mathrm{L} = \sum_{c=1}^{n_c} y_c^{ref} \log P(y_c^{ref} | x^{ref}) + \lambda \|W\|_2^2$$

where λ = 0.01 and controls weight decay to prevent overfitting and y_c^{ref} are reference cell type labels. Gradient updates use:

$$W_{t+1} = W_t - \eta t \frac{\partial \mathrm{Lref}}{\partial \mathrm{W}} \quad \text{(frozen for query prediction)}$$

where W_t is the model weights at training step t, ηt is the learning rate, and $\frac{\partial \mathrm{Lref}}{\partial \mathrm{W}}$ is the gradient of reference-only loss.

The model is optimized via AdamW with learning rate decay. Gradient norms are clipped to 1.0 to prevent explosion.

Query Prediction Protocol

Unlabeled query cells are processed through the frozen model:

$$P\left(y_c^{ref} \middle| x\right) = softmax\left(f_{\theta_{frozen}}(x_{query}^{pathways})\right)$$

Composite Score Calculation

For each cell, the model computes attention-weighted pathway activity scores, resulting in a matrix $A \in \mathbb{R}^{N*p}$, where N is the number of cells and p is the number of pathways. For each gene of interest, the model provides a vector of attention weights over all pathways from a learned matrix $P \in \mathbb{R}^{N*p}$. This matrix encodes the relevance of each gene to each pathway, as learned through training. For a set of target genes, their corresponding rows from P were extracted, yielding a weighted matrix $W \in \mathbb{R}^{N*k}$. To compute the composite scores, we project the pathway activity of each cell onto the gene-specific weight vectors. This matrix

multiplication yields a composite score per cell, reflecting the aggregated influence of each gene across latent pathways. These composite scores were subsequently stored for visualization and downstream analysis.

Visualization Framework

i. Heatmap

Pathway activation scores are z-scored across cells for heatmap visualization:

$$A_{z-score} = \frac{A_{i,j} - \mu_j}{\sigma_j}$$

where μ_j and σ_j are the mean and standard deviation of pathway *j*'s activity.

ii. UMAP

UMAP reduces 512-dimensional pathway embeddings to two dimensions using cosine distance and min_dist = 0.3:

$$Z_{UMAP} = UMAP(PCA(A_{z-score}, n_{comps} = 50), \text{min\_}dist = 0.3)$$

The first UMAP is colored by their annotated cell type labels, and a second UMAP displays cells colored according to composite scores calculated for each target gene.

iii. Violin Plot

For each gene, composite scores were compared between disease and control groups using violin plots.

Implementation Details

QuokkaVision was developed using PyTorch 2.2.2 and Scanpy 1.10.4, with 8 parallel heads, a 0.2 dropout rate, and trained for 15 epochs (batch size=256)

2.3 Validation of QuokkaVision

Full Model ROC Validation

The full QuokkaVision model (with pathway mask) was evaluated on held-out reference cell-type data using one-vs-rest receiver operating characteristic (ROC) analysis (**Figure 5**). Cell type labels $y \in \{1, ..., n_c\}$ were binarized into $y_{bin} \in \{0, 1\}^{m \cdot n_c}$, and AUC scores computed for each class.

Ablation Study for Pathway Contribution

To quantify the impact of pathway integration, two model variants were compared: (1) full QuokkaVision architecture incorporating the pathway-gene matrix *M*, and (2) an ablated variant where *M* was replaced by a zero matrix (**Figure S4**). Both models were trained on reference data using identical hyperparameters. Statistical significance of differential AUC

values was assessed using a paired Wilcoxon signed-rank test, with the test statistic W defined as:

$$W = \sum_{c=1}^{n_c} rank(|\Delta AUC_c) \cdot sign(\Delta AUC_c),$$

where rank denotes the absolute value ranking of ΔAUC_c.

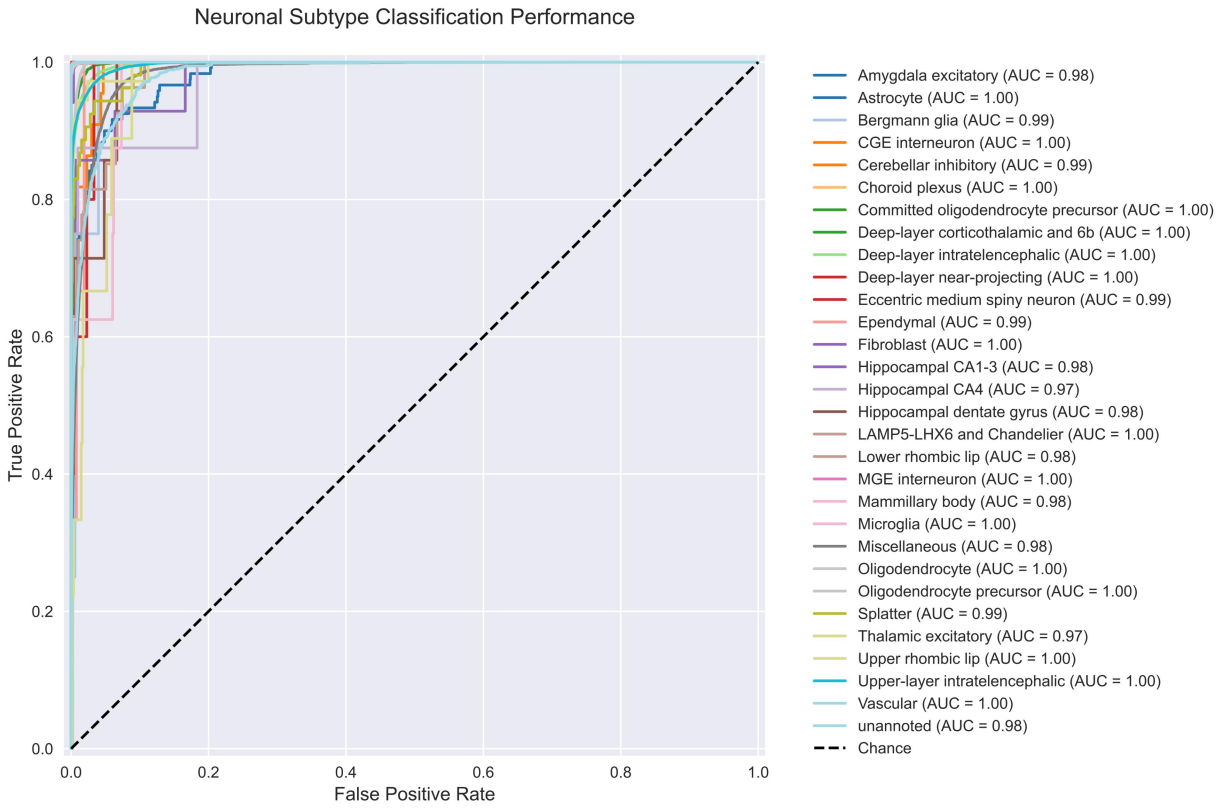


Figure 5. ROC curves of prediction accuracy by neuronal subtype. The black dashed line indicates chance performance (AUC = 0.5). Curves are annotated with per-class AUC values, rounded to two decimal places.

2.4 Functional and In Vivo Validation of Candidate Genes in C. elegans

C. elegans culture and maintenance

Two *C. elegans* strains were used in this study: GRU102 and CL2006. GRU102 expresses pan-neuronal truncated human Aβ ($A\beta_{1\text{-}42}$) under the control of the *Pmyo-2* promoter and exhibits a progressive neurodegenerative phenotype, including cognitive deficits such as impaired associative memory (Fong et al., 2016). In contrast, CL2006 expresses $A\beta_{1\text{-}42}$ in body wall muscle cells under the control of the *Punc-54* promoter, which enabled it to present a progressive, adult-onset paralysis and premature death (Lublin & Link, 2013). GRU102 was

used exclusively in body length, speed, and chemotaxis assays to assess overall health, neuronal function, and memory, while CL2006 was used in RNA interference (RNAi) and Aβ deposition assays.

Both strains were maintained on nematode growth medium (NGM) agar plates seeded with *Escherichia coli* OP50 at 20°C. Synchronized populations were obtained by 2-hour egg-laying at 20°C, then cultured to adulthood for 6 days. Adults were transferred to NGM plates with 5-fluoro-2'-deoxyuridine (FUdR) treatment.

Stains and Crosses

Double-stranded RNA interference (dsRNAi) was performed on the CL2006 strain. HT115(DE3) bacteria carrying either the empty vector pL4440 (negative control) or double-stranded RNA (dsRNA) targeting *Htt-1*, ortholog of *Htt*, were cultured in lysogeny broth (LB) containing carbenicillin (25 μg/mL), tetracycline (15 μg/mL), and isopropyl β-D-1-thiogalactopyranoside (IPTG; 5 mg/mL) for 6 hours at 37°C. Synchronized L4-stage CL2006 worms (n = 30 per group, 3 biological replicates) were transferred to RNAi plates seeded with HT115 bacteria expressing *htt-1* dsRNA or pL4440 controls.

Dose-response Toxicology Assays

Dose-dependent toxicity profiles of candidate drugs were established in Aβ-expressing CL2006 worms. CL2006 strains were transferred to NGM plates with test compounds (0.1% DMSO vehicle, 0.1 μM SAHA, 10 μM SAHA, 0.6 mM VPA, or 1.2 mM VPA). Survival was quantified after 144 hours.

Pharmacological Treatments

To evaluate the potential therapeutic efficiency targeting *Htt*, two HDAC inhibitors were selected, namely suberoylanilide hydroxamic acid (SAHA; Sigma-Aldrich) and valproic acid (VPA; Selleck Chemicals LLC). Drug solutions were prepared by dissolving 10 μM SAHA or 0.6 mM VPA in 0.1% dimethyl sulfoxide (DMSO) and incorporated into NGM agar during plate preparation. Synchronized L4-stage CL2006 worms were exposed to drug-supplemented plates or vehicle control for 7 days at 20°C. Experimental cohorts include: (1) 10 μM SAHA, (2) 0.6 mM VPA, (3) combination treatment: 10 μM SAHA and 0.6 mM VPA, and (4) vehicle control (n = 50 per group, 3 biological replicates.

Lifespan Assays

All lifespan assays were conducted using synchronized L4-stage of CL2006 worms (n = 50 per group; 3 biological replicates) maintained at 20°C on freshly prepared drug or control plates. To prevent progeny contamination, worms were transferred daily to fresh NGM plates during the egg-laying period (days 1-3 of adulthood) and every 3-4 days thereafter. For each cohort of worms, survival was assessed daily; worms were scored as dead if unresponsive to

gentle stimulation. Animals that crawled off plates, exhibited internal hatching, or displayed a ruptured vulva were censored.

Body Length and Speed Quantification

Locomotor performances were quantified in L4-stage CL2006 and GRU102 worms (n = 20 per group) acclimated for 5 minutes on fresh NGM plates without OP50 bacteria (**Figure 6**). Body length and speed were quantified via WormLab under a stereomicroscope over a 30-second interval per worm (MBF Bioscience LLC, Williston, VT USA).

Chemotaxis Assay

Chemotaxis assay was performed on GRU102 worms (**Figure 6**). Odorant preference was assessed using a 6 cm NGM plate divided into quadrants. The attractant (0.1% benzaldehyde) and control (100% ethanol) solutions were supplemented with 1 µL 1 M sodium azide and applied to opposing quadrants. Synchronized adult worms (n = 30 per group; 3 biological replicates) were placed at the plate center and allowed to navigate for 60 minutes at 20°C. The chemotaxis index (CI) was calculated as:

$$CI = \frac{\text{number of worms at attractant location} - \text{number of worms at control location}}{\text{total number of worms on the plate}}$$

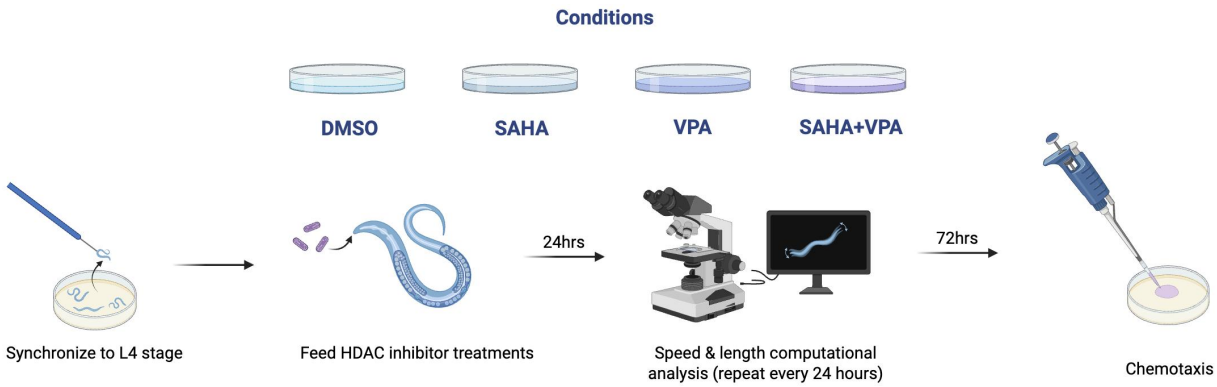


Figure 6. Workflow for *C. elegans* Motor and Chemotaxis Assays (created with Biorender).

Aβ Deposition Assay

Aβ deposition assay was conducted with L4-stage CL2006 (n = 30 per group). Worms were fixed in 4% paraformaldehyde/PBS (pH 7.4) for 24 hours at 4°C, followed by permealization in 5% fresh β-mercaptoethanol, 1% Triton X-100, 125 mM Tris (pH 7.4) at 37 °C for 24 hours. Worms were stained with 0.125% thioflavin S (Sigma-Aldrich) in 50% ethanol for 2 min, destained in 50% ethanol for 2 minutes, washed with PBS and mounted on slides prepared with agar pads for visualization. Fluorescence imaging was carried out using a Nikon SMZ18 fluorescence microscope.

Statistical Analysis

Lifespan data were analyzed using Kaplan-Meier survival curves with log-rank tests. Body length and speed were compared across groups via two-way ANOVA followed by Tukey's honestly significant difference (HSD) post-hoc test. Chemotaxis indices were compared by one-way ANOVA with pairwise post-hoc test.

3. Results

3.1 Integrated PPI Network Reveals a Htt-HDAC Axis in MDD and AD

To investigate the molecular interactions between MDD and AD, we overlapped the DEGs of two independent MDD GWAS datasets and constructed a protein-protein interaction (PPI) (**Figure 7**). Through the PPI analysis, we revealed a layered genetic architecture. Notably, a highly interconnected network of canonical histone variant genes, including *H2BC15*, *H2BC17*, *H4C8*, *H2AC13*, *H3C12*, *H1-5*, *H2BC13*, interacts with *Htt* and *CACNA2D1* (**Figure 7**). This histone-encoding network is particularly significant given that HDACs have been identified as key pathophysiological markers of MDD, and six out of seven histone-related genes in this network are involved in HDAC-mediated histone deacetylation pathways (**Figure S5**).

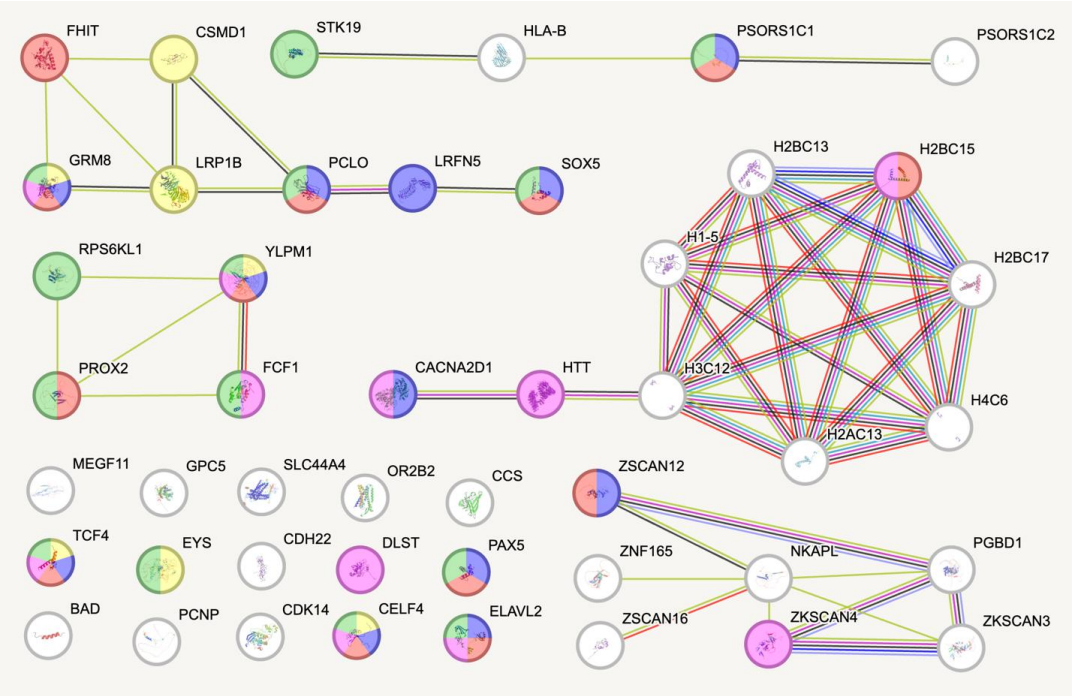


Figure 7. Protein-Protein Interaction (PPI) Network of DEGs from Two MDD GWAS Datasets. Nodes were color-coded based on psychological measurement layers: depressive symptom measurement (red), wellbeing measurement (blue), neuroticism measurement (green), mood instability measurement (yellow), and emotional symptom measurement (pink).

Intriguingly, genes uncommonly associated with MDD or AD are linked to histone-related genes, such as *Htt*, which is commonly associated with Huntington's disease (HD), and *CACNA2D1* (Calcium Voltage-Gated Channel Auxiliary Subunit Alpha2delta 1), which is often linked to early-onset developmental epileptic encephalopathy (**Figure 7**). Their striking presence suggests a potential role as a broader link between MDD and AD.

The functional coherence of these interactions was also confirmed by pathway enrichment analysis (**Figure 8**). MDD-associated DEGs from overlapped GWAS datasets were enriched for regulation of cellular homeostasis (FDR = 1.2e−7), generation of neurons (FDR = 4.1e−7), and regulation of axonogenesis (FDR = 4.7e-6), processes central to both disorders. The main enriched pathways identified through GWAS data comparing MDD patients and AD patients coincide with the role of *Htt* in axonogenesis, neural development, neuron differentiation, and its indirect role in the regulation of cell size. This positions the *Htt*-HDAC network as a network participating in MDD pathology.

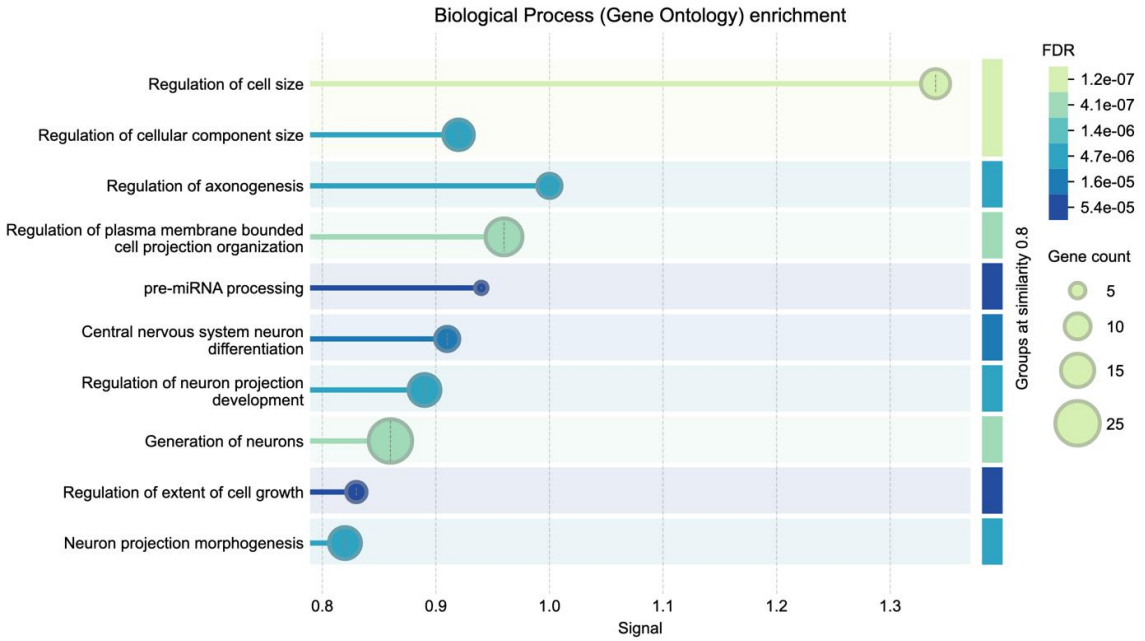


Figure 8. Biological Process Enrichment of DEGs Between MDD and Control Groups. The signal is graphed with respect to the false discovery rate (FDR).

3.2 Htt as a Transcriptional Hub Linking Epigenetics and Synaptic Dysfunction

Next, to reveal the downstream effects of the *Htt*-HDAC network, we constructed a regulatory network with *Htt* as a central mediator. This network revealed interactions between *Htt* and genes such as *APBB2*, *ELAVL3*, and *PNISR*, which are genes known to involve in neuronal function synaptic integrity and linked to both AD and MDD (**Figure 9; S7**). For example, *ARX* upregulation (identified in transcriptomic analysis) suggests *Htt*'s involvement in RNA regulation, potentially via its interaction with *ELAVL3*. This interaction could stabilize ARX transcripts, amplifying inhibitory signaling and destabilizing the cortical excitation/inhibition

balance. Similarly, other *Htt*-associated genes, including *APBB2* (linked to increased risk of late-onset AD and depression) and *PNISR* (promising candidate for AD), further support the network's relevance to AD and MDD (**Figure 9**). As implicated from earlier pathway enrichment with GWAS data, loss of *Htt* may also cause axonal transport defects and disrupt the transportation of BDNF, which is also a pathological event observed in AD and MDD (Gauthier et al., 2004). These downstream genes are also analyzed for their expression level across six cell types in MDD patients (**Figure S8**). In short, the *Htt*-HDAC network emerges as a promising link between AD and MDD, mediated by synaptic signaling, axonal transport, and direct regulatory mechanisms.

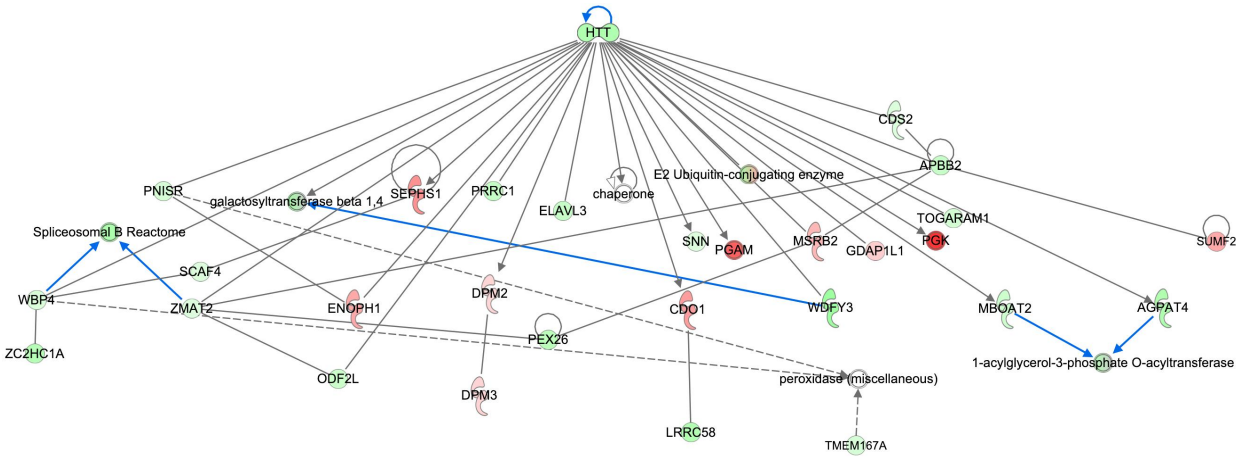


Figure 9. *Htt* as an Upstream Regulator of Potential Pathways Linking AD and MDD. Previous *Htt* regulatory network also positions *Htt* as a central hub in regulating neurodevelopment (Poplawski et al., 2020). Here, *Htt* is associated with other key genes in the disease progression of AD and MDD, including *PNISR*, *APBB2*, and *ELAVL3*).

3.3 Transcriptomic Profiling Highlights Neurodevelopmental, GABAergic, Histone Dysregulation Across Subtypes

To explore these interactions on a transcriptomic level, we performed transcriptomic profiling in search of DEGs between MDD patients and healthy controls (**Figure 10a**). This revealed significant downregulation of *NNAT* (responsible for cellular proliferation), *NKX2-2* (involved in embryonic patterning), and *NLRP2* (an inflammasome component involved in neurodevelopmental processes), as well as the upregulation of *ARX* (a regulator of GABAergic neuron differentiation). These findings are consistent with early AD pathology for which the disrupted neurogenesis and GABAergic dysfunction precede amyloid accumulation.

Building on this, we next performed neuron-specific DEG analysis to further explore the molecular pathology of MDD (**Figure 10b-e**). Expected overlaps between neuron-specific DEGs and pan-cellular neuronal signatures are observed (*NNAT*, *NKX2-2*, *ARX*). Interestingly, we observed the consistent dysregulation of histone-encoding genes (*HIST1H2AC*, *H3F3B*, *H1F0*, and *H1FX*) across five out of six neuronal subpopulations (**Figure 10b-e; S6**). These

genes are also part of the *Htt*-HDAC network. The recurrent dysregulation of histone-related genes also validates the role of chromatin-level regulation in MDD pathophysiology and the significance of the *Htt*-HDAC network.

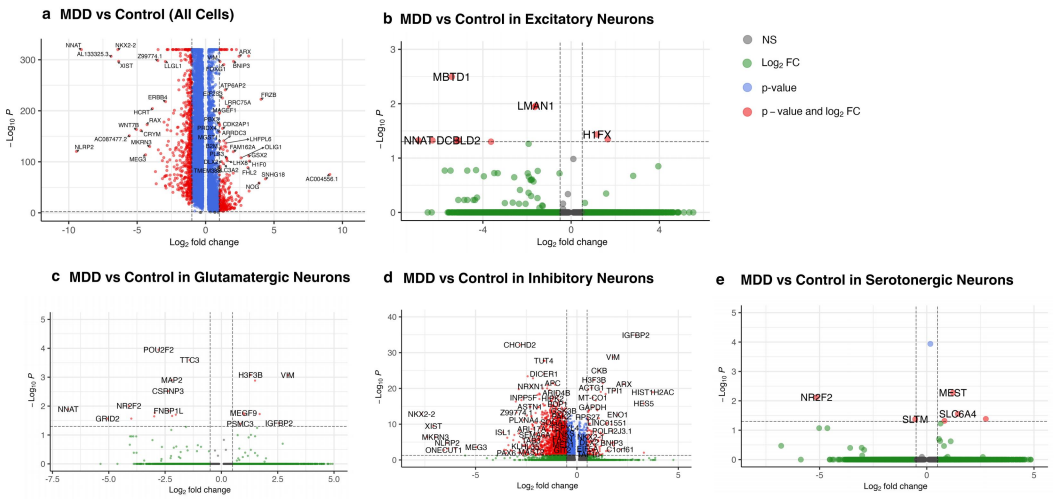


Figure 10. Transcriptomic Analysis of DEGs in MDD and Control Groups. For (a) all neurons, genes were considered significantly differentially expressed with a p-value < 0.01 and $|\log_2 FC| \geq 1.0$. For (b) excitatory neurons, (c) glutamatergic neurons, (d) inhibitory neurons, and (e) serotonergic neurons, genes were considered significantly differentially expressed with a p-value < 0.05 and $|\log_2 FC| \geq 0.5$. Note that glutamatergic neurons are a subtype of excitatory neurons, and it is isolated from the rest of the excitatory neuron group for the analysis in (c).

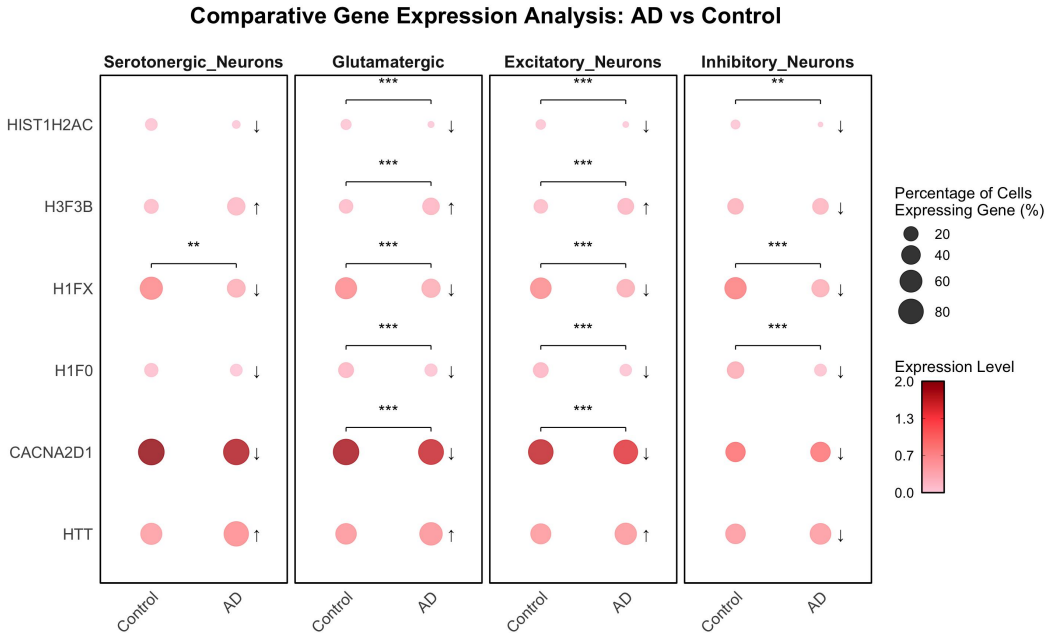


Figure 11. Transcriptomic Analysis of MDD DEGs and Histone-related Genes in AD and Control Groups. Asterisks indicate statistical significance: p < 0.01 (**), and p < 0.001 (***).

To validate the significance of our network in AD, we mapped DEGs and histone-related genes identified in MDD onto AD-specific scRNA-seq datasets. Notably, excitatory neurons in AD patients exhibited significant expression differences in *CACNA2D1*, *HIST1H2AC*,

H3F3B, *H1FX*, and *H1F0* compared to controls, with insignificant upregulated expression level shifts for *Htt* (**Figure 11**). This confirms the genetic overlaps between AD and MDD and the centrality of histone-related genes in both diseases; while the insignificance of *Htt*'s expression changes suggests that *Htt* is not regulated on a transcriptional level compared to the other genes in the network. Given the close interaction of *Htt* with histone variants and the lack of significant transcriptomic changes, we hypothesize that its regulatory role operates at the chromatin structural level rather than through transcription.

3.4 Chromatin Accessibility Shifts at the Htt Locus in AD Neurons

To test our hypothesis that *Htt* functions at the chromatin level, we revealed chromatin-level modifications at the *Htt* locus undetectable at the transcriptional level. Using the UCSC Genome Browser, we analyzed chromatin accessibility in two neuronal subtypes previously implicated in MDD pathway analyses from QuokkaVision: SST inhibitory neurons and L5 excitatory neurons (**Figure 12**).

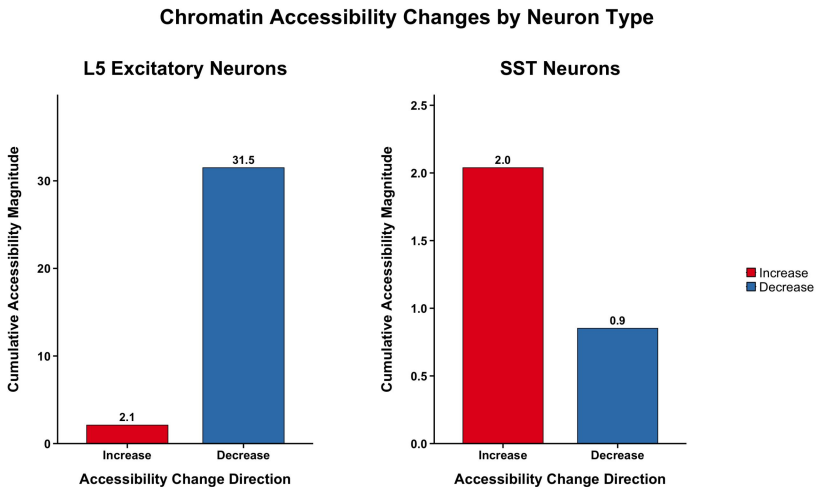


Figure 12. Differential Chromatin Accessibility for L5 Excitatory Neurons and SST inhibitory Neurons from AD patients (ADNC3 vs. Control).

In late-stage AD (NC3), SST inhibitory neurons exhibited noticeably increased chromatin accessibility at the *Htt* locus compared to controls (NC0) (**Figure 12**). Conversely, L5 excitatory neurons showed a markedly decrease in accessibility at the same locus (**Figure 12**). These divergent epigenetic profiles hold critical importance as excitatory-inhibitory (E/I) imbalance was identified as one of the important pathological markers of both diseases. These analyses position *Htt* as a chromatin scaffold regulating epigenetic modifications.

In diseased L5 excitatory neurons, reduced accessibility at the *Htt* locus disrupts its scaffolding function. Reduced chromatin openness limits *Htt*'s ability to recruit HDACs or architectural proteins, which is evident in the impaired calcium signaling (*CACNA2D1* downregulation) and axonal transport (*FEZF2*/*PDYN* loss). This loss of chromatin plasticity exacerbates excitotoxicity and synaptic instability, driving neurodegeneration. While in L5-6 SST inhibitory neurons, increased accessibility at the *Htt* locus enables compensatory

chromatin remodeling. Here, *Htt* acts as a docking site for HDACs, which repress SST enhancers while activating stress-response genes (such as *TH*) through chromatin looping, where *Htt*'s open chromatin physically interacts with distal regulatory elements. This spatial reorganization allows stress-responsive transcription factors to bind and upregulate compensatory pathways, temporarily preserving inhibitory pathway activation despite global SST repression. This suggests that the *Htt*-HDAC network drives neurodegeneration as well as potential compensatory mechanisms through chromatin plasticity.

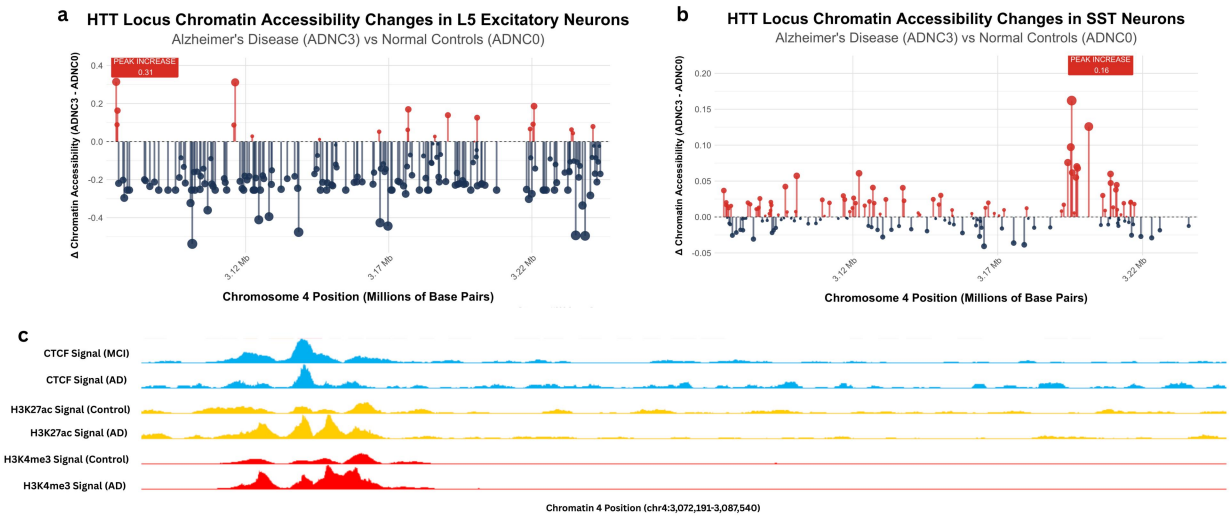


Figure 13. Chromatin Accessibility Changes at the *Htt* Locus in AD patients (ADNC3 vs. Control). Changes are mapped according to chromosomal position in (a) L5 excitatory neurons and (b) SST inhibitory neurons. (c) H3K27ac and H3K4me3 signals in healthy control and AD patients, and CTCF signals in MCI and AD patients are mapped (CTCF data for healthy controls is unavailable).

Further analyses of H3K27ac and H3K4me3 signals between healthy controls and AD patients revealed a slight increase in H3K27ac and an approximate two-fold rise in H3K4me3 activity at the *Htt* promoter region, consistent with our discovery of selective increased accessibility in specific neuronal populations, despite overall reduced accessibility (**Figure 13c**). Together, these findings suggest that the *Htt*-HDAC network operates through dynamic chromatin accessibility, with histone-related genes modulating *Htt* as upstream regulators, and *Htt*, in turn, influencing downstream disease-relevant pathways.

3.5 QuokkaVision Links Htt-HDAC Network to Layer-Specific Dysregulation

To map the established *Htt*-HDAC network onto genetic and molecular pathways while preserving biological nuanced data, we developed a transformer model, QuokkaVision, to highlight the regulation of pathways in relation to the identified genes. QuokkaVision can be trained on multiple different pathway datasets to map neuronal subtypes and genetic markers to specific activated or repressed pathways. Here, we have QuokkaVision trained on the Allen Brain Atlas (genetic), KEGG Human 2021 (metabolic), and Reactome Pathways 2024 (signaling) to analyze MDD scRNA-seq datasets.

As expected, QuokkaVision highlighted pathways that are highly correlated with excitatory and inhibitory neurons dysregulation (**Figure 14**). Notably, there's a pronounced increase in activation in the Excitatory L5 *FEZF2 NREP-AS1* and Excitatory L5-6 *FEZF2 IFNG-AS* pathways. *FEZF2* is implicated to be associated with the neurodevelopmental phenotypes in recent studies (Garber et al., 2024). At the same time, inhibitory L5 neurons are also associated with pathways that are uniformly downregulated in cerebellar inhibitory, hippocampal CA1-4, and thalamic excitatory neurons (**Figure 14**). This centers L5 excitatory and inhibitory neurons as important candidates for the detection of pathway alterations in MDD.

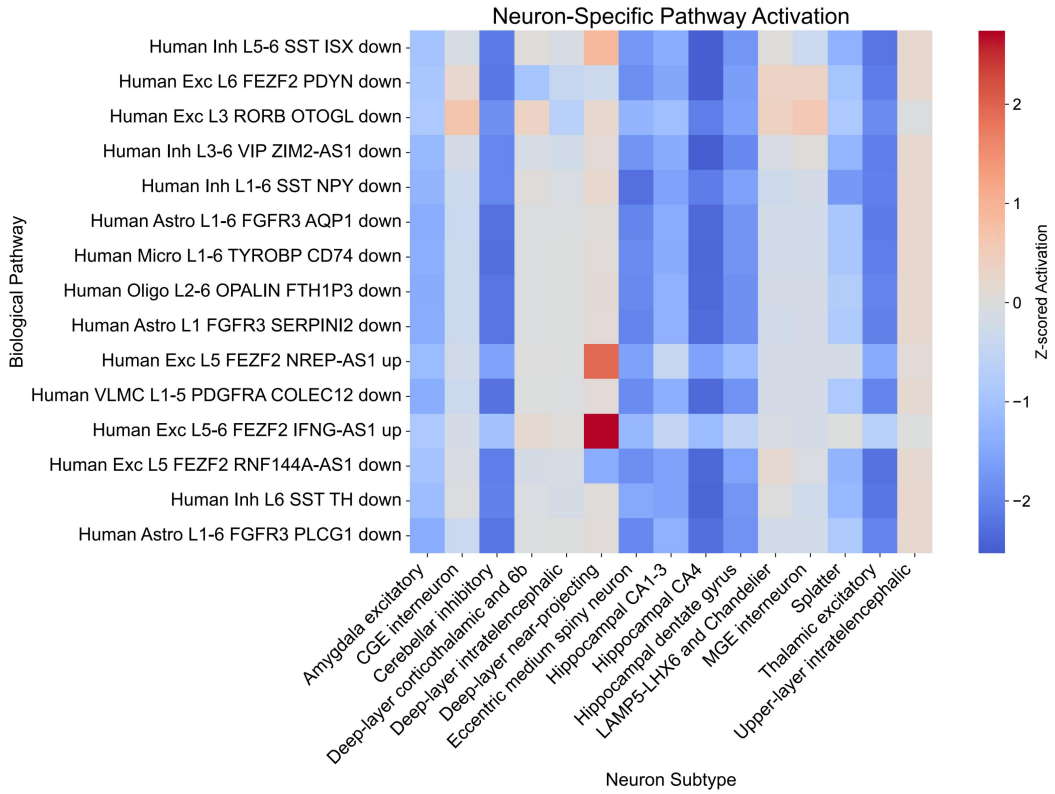


Figure 14. Neuron-specific Genetic Pathway Activation in MDD Patients Mapped Against Allen Brain Atlas Pathway Mask. This analysis is generated by QuokkaVision, an AI-powered, transformer-based LLM model providing a sophisticated computational framework for high-precision mapping.

To reveal specific pathways associated with the *Htt*-HDAC network, QuokkaVision systematically mapped *Htt*, *HIST1H2AC*, *CACNA2D1*, *H1F0*, *H1FX*, and *H3F3B* onto their respective pathways (**Figure 15**). QuokkaVision highlighted a uniform downregulation of inhibitory neurotransmission pathways, including SST (Somatostatin) and NPY (Neuropeptide Y), across L1-6 inhibitory neurons, consistent with transcriptomic deficits in GABAergic signaling, which plays a significant role in both AD and MDD (**Figure 14**). This aligns with prior studies where mRNA levels of SST were found to decrease in brain with AD. However, in a striking divergence, histone-related genes (*H1F0*, *H1FX*, and *H3F3B*) were associated with upregulated activity in the L5-6 SST pathway, despite global suppression of SST

expression (**Figure 14 & 15**). This anti-intuitive co-occurrence of pathway activation and gene repression potentially suggests a compensatory mechanism: while SST transcription is universally reduced, in some critical populations, epigenetic modification at specific loci like *Htt* may amplify residual SST signaling. The *Htt*-HDAC network, which integrates histone modifiers and calcium channels (*CACNA2D1*), emerges as a likely mediator of this compensation.

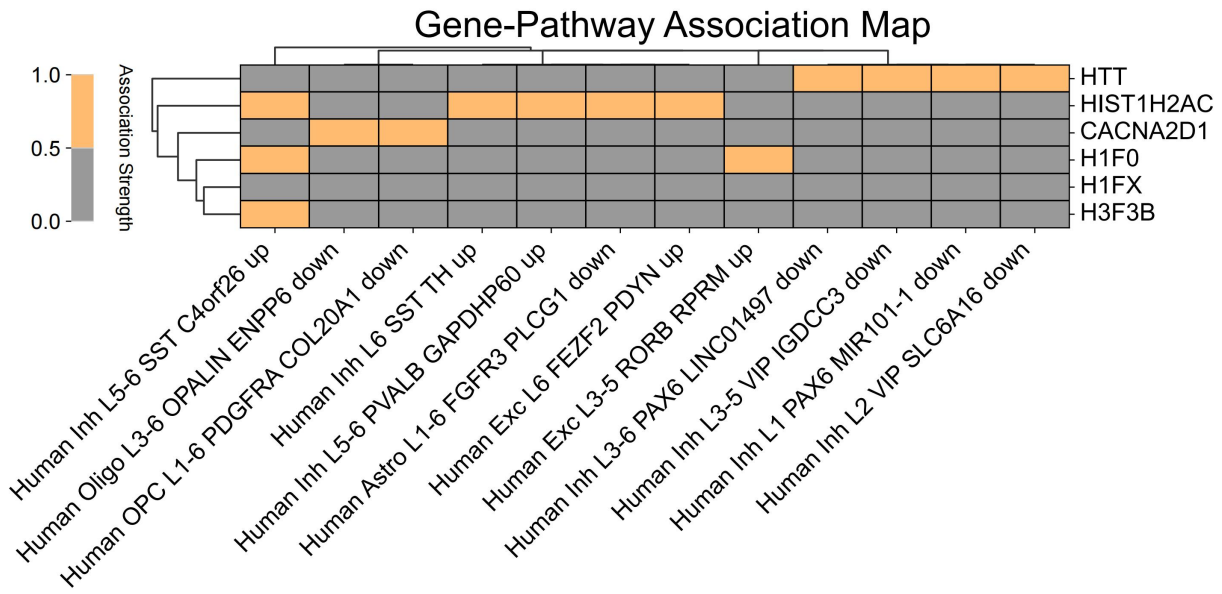


Figure 15. Gene-to-pathway Association Map of the *Htt*-HDAC Network. This analysis is generated by QuokkaVision, an AI-powered, transformer-based LLM model providing a sophisticated computational framework for high-precision mapping.

To further validate the significance of the *Htt*-HDAC network, we conducted analyses with KEGG and Reactome Pathway Masks. Analyzing all neuronal subtypes, QuokkaVision revealed enrichment in the following pathways: chromatin organization, epigenetic regulation of gene expression, and chromatin-modifying enzymes (such as HDACs and histone acetyltransferases) pathways (**Figure S9**). It also confirms the relevance of previously established pathways, such as serotonergic synapse, long-term depression, MAPK signaling, ERBB signaling, etc., in depression (**Figure S10**) (Blier & El Mansari, 2013; Chakraborty et al., 2019; Chen et al., 2022; Du et al., 2019; Mango et al., 2019; Zhou et al., 2021). These findings solidified the role of epigenetic modulation in regulating MDD and identify *Htt* as an epigenetic modulator that influences downstream targets, thereby contributing to AD- and MDD-related phenotypes.

3.6 Functional Validation of the Htt-HDAC Network via dsRNAi Suppression

To establish roles of the *Htt*-HDAC in vivo, we targeted *htt-1* (ortholog of human *Htt*) in $A\beta_{1-42}$-expressing CL2006 *C. elegans* using dsRNAi. Synchronized L4-stage worms (n = 50 per group, 3 biological replicates) exposed to *htt-1* dsRNA exhibited a significant delay in paralysis onset and an approximately 4-day extension in lifespan ($p < 0.001$) (**Figure 16**).

These findings indicate that, loss of *htt-1* yields a survival advantage under Aβ-induced stress. This observation aligns with the transcriptional expression analysis, which revealed that *Htt* is upregulated in AD patients (**Figure 11**), raising the possibility that *Htt* may promote Aβ-related pathways that accelerate neurodegeneration and reduce lifespan. To test this hypothesis, thioflavin S staining was performed in CL2006 worms (n = 20 per group, 3 biological replicates). Quantitative analysis revealed a significant reduction in integrated Aβ fluorescence intensity in *htt-1* RNAi worms compared with controls (L4440 and HT115). This demonstrates that *htt-1* knockdown reduces Aβ accumulation in vivo, supporting our hypothesis that *Htt* actively contributes to Aβ deposition.

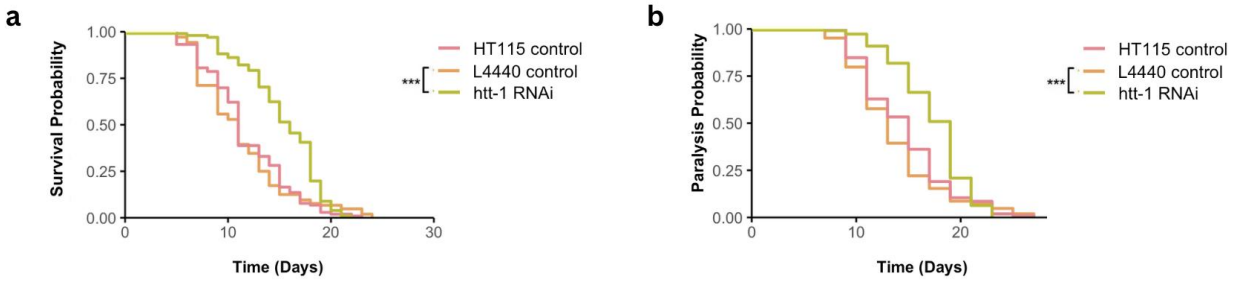


Figure 16. *htt-1* RNAi Knockdown Effects on Paralysis and Survival Assays. RNAi knockdown (a) delays paralysis onset by 4 days, and (b) extends lifespan by 4 days. Asterisks indicate statistical significance: $p < 0.001$ (***).

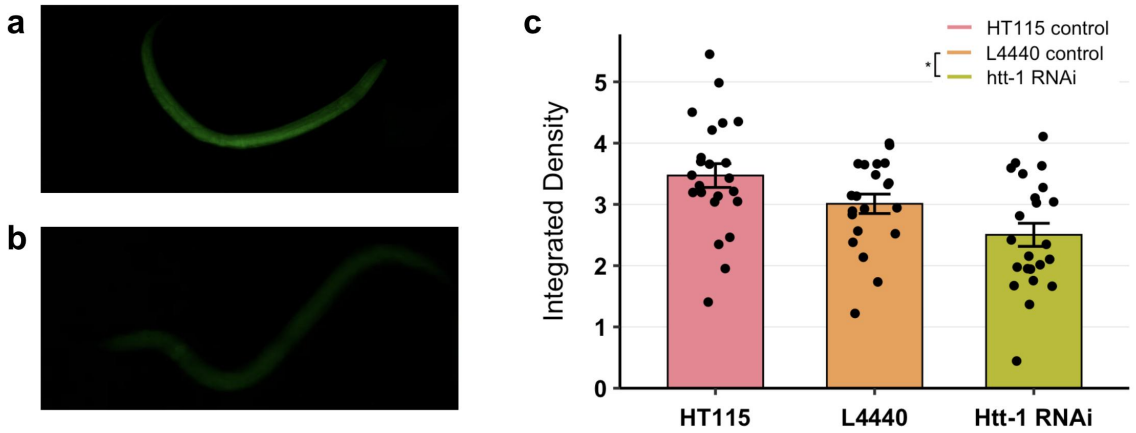


Figure 17. *htt-1* RNAi Knockdown Effects on Aβ deposits. Representative fluorescence images of CL2006 worms: (a) control (no RNAi) and (b) *htt-1* RNAi knockdown. (c) Integrated intensity of fluorescence signal across three groups: *htt-1* RNAi, L4440, and HT115 (calculated as area × mean fluorescence intensity). Asterisks indicate statistical significance: $p < 0.05$ (*).

3.7 Therapeutic Efficacy Evaluation via Pathological Inhibition of the Htt-HDAC Network by HDAC Inhibitors

Building on genetic validation, we pharmacologically targeted the network using HDAC inhibitors to determine therapeutic efficiency. To rigorously establish the optimal drug concentrations for *C. elegans*, dose-response toxicology experiments were conducted with SAHA at 0.1 and 10 μM and VPA at 0.6 mM and 1.2 mM (**Figure S11**). Candidate concentrations were selected based on prior studies establishing non-toxic ranges (Chen et al.,

1999; Grigolon et al., 2022). Survival assays confirmed 100% viability of *C. elegans* at 10 µM SAHA and 0.6 mM VPA individually (**Figure S11**). These concentrations were thus applied in drug treatment.

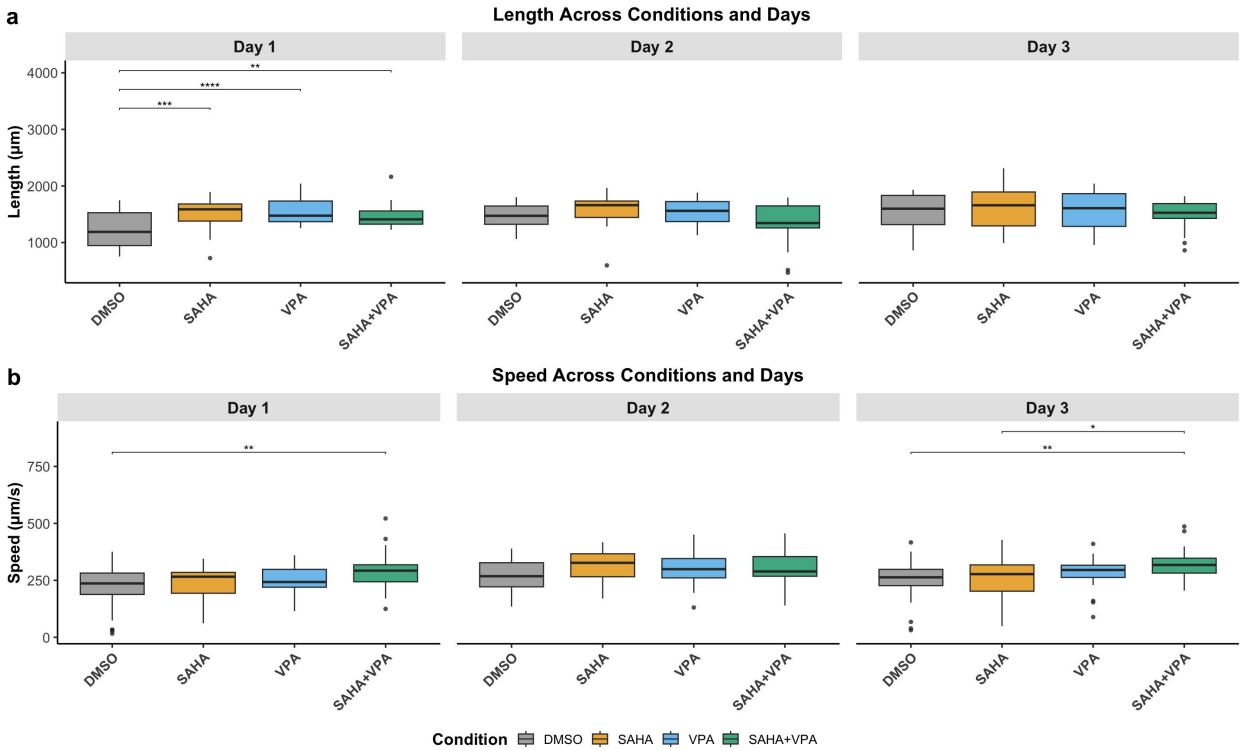


Figure 18. Length and Speed Measurements of GRU102. Measurements of GRU102 *C. elegans* were taken under different HDAC inhibitor treatments over three days following the L4 stage (n = 3 biological replicates; n = 20 per group). (a) Mean body length (µm) and (b) mean locomotion speed (µm/s) of worms treated with SAHA, VPA, SAHA+VPA combination, or DMSO vehicle control. Asterisks indicate statistical significance: $p < 0.05$ (*), $p < 0.01$ (**), $p < 0.001$ (***), and $p < 0.0001$ (****).

To further confirm overall health and cognitive function, the GRU102 worm model was adopted. Worms were fed with HDAC inhibitors at the optimal concentrations, and worm length, speed, and chemotaxis index were measured post-treatment day 1, 2, and 3 (**Figure 18, 19**). Dual HDAC inhibition with SAHA+VPA produced an 18.7% increase in body length on day 1 (1466 µm vs DMSO: 1235 µm, $p < 0.01$) (**Figure 18**). Locomotor-wise, SAHA+VPA treated worms achieved speeds of 324 µm/s, which is 27.5% faster than DMSO (254 µm/s, $p < 0.01$) (**Figure 18**). Additionally, chemotaxis assays revealed the combination therapy's cognitive enhancement: SAHA+VPA (CI = 0.551) generated a 39.5% greater chemotaxis index than SAHA monotherapy (CI = 0.395) and 224% improvement over VPA alone (CI = 0.170, $p < 0.001$), while completely reversing the negative chemotaxis observed in DMSO controls (CI = - 0.504, $p < 0.001$) (**Figure 19**). Collectively, these data demonstrate a potential synergistic effect with SAHA and VPA in the GRU102 model, significantly surpassing VPA monotherapy effects (**Figure 19**).

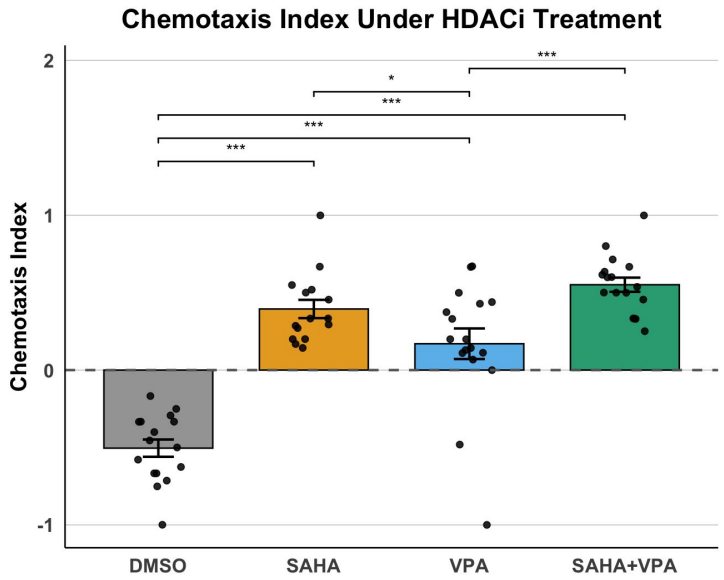


Figure 19. Chemotaxis index of GRU102 under HDAC inhibitor treatments (n = 3 biological replicates; n = 50 worms/group; 5 technical replicates of 10 worms each). The dashed horizontal line at zero indicates no chemotaxis preference. Asterisks indicate statistical significance: $p < 0.05$ (*) and $p < 0.001$ (***).

4. Discussion

Although emotional and cognitive changes in aging are increasingly recognized as connected, research and clinical practice often treat them separately. Cognitive changes, for example, are typically linked to overall brain structure, while emotional shifts are associated with specific neural circuits (Knight et al., 2020; Mather, 2024). In older adults, this is especially concerning: symptoms like forgetfulness or mood changes are often broadly labeled as "normal aging," even though the connection between them remains unclear. This oversimplification can hide early pathological changes, delaying early intervention for neurodegenerative or neuropsychiatric conditions. Critically, the shared biological mechanisms behind these co-occurring symptoms remain poorly understood. Here, we address this gap by identifying the *Htt*-HDAC axis as one of such important pathways linking AD and MDD. Through in vivo validation, we show that this axis affects both lifespan and cognitive performance, providing mechanistic insight into the connection between emotional and cognitive changes in aging. Recognizing this shared pathway holds significant potential for developing improved clinical strategies to better serve aging populations.

4.1 QuokkaVision in Context: From Architecture to Implications

Here, we developed QuokkaVision, which is a transformer-based LLM framework that directly links molecular pathways (related to neurodegeneration and neuropsychiatric disorders) to quantifiable genetic dysregulation in specific neuronal populations. Unlike conventional gene-centric models, which often prioritize differential expression of genetic markers without mechanistic context (Adema et al., 2024; Liu & Trapnell, 2016),

QuokkaVision embeds pathway structure into the model architecture and supports gene-to-pathway attribution under a cell-type-specific resolution. Building on prior biological knowledge-integrated networks (Chen et al., 2023), and recent LLMs for transcriptomic analyses (Brixi et al., 2025), our model reveals region-specific circuits, such as heme signaling disruption, ERBB2 and ERBB4 signaling, and chromatin instability, that may be obscured in gene-level analyses. With multi-head biological attention, QuokkaVision maps functional pathway activity at single-cell resolution, bridges molecular identity with functional dysregulation, and supports hypothesis-free, region- and subtype-specific discovery. Specifically, in this study, QuokkaVision enabled us to reveal the scaffolding role of *Htt* in neurodegenerative and neuropsychiatric disorders.

Intriguing, QuokkaVision unbiasedly implicated that different cortical layers exhibit divergent chromatin responses to cognitive impairment: one epigenetic regulatory network may be compensatory in some layers, while pathological in others. This raises a fundamental question: *To what extent do layer-specific chromatin states dictate disease trajectories*? Based on our hypothesis, we propose a concept of "spatial epigenetic adaptation" to describe circuit-specific chromatin regulation. In this model, cognitive impairment drives epigenetic changes on a cellular level, triggering different responses across cortical layers. For *Htt*, expression is downregulated in most layers but upregulated in some, potentially preserving its other essential functions (**Figure S12**). This adaptative characteristic allows neuronal populations to modulate gene expression in a layer-specific way that responds to circuit-specific demands or pathological stress. Here, *Htt* acts as a critical mediator to recruit chromatin modifiers to form compensatory loops that delay synaptic plasticity loss and neurodegeneration.

4.2 Implications of the Htt-HDAC Axis

The *Htt*-HDAC axis centers on *Htt*, which is implicated to be a gene regulating lifespan, motor control, and cognition. A previously published study on neural progenitor cells (NPC) highlights *Htt*'s key role in regulating neural plasticity and restoring gene activity patterns (Poplawski et al., 2020). In mice, *Htt* deletion at embryonic stage E12 significantly impairs plasticity regeneration (Poplawski et al., 2020). Interestingly, this goes against our conclusion of the detrimental effects of *Htt*. However, note that our study selected samples in their adult or elderly stages. This immediately points to age-related factors: *Htt* may have temporal-dependent roles. During embryonic and developmental stages, *Htt* builds and restores neural plasticity, and in the late years, potential epigenetic alterations in *Htt* contribute to shortened lifespan and impaired motor and cognitive abilities. These epigenetic alterations, as we see, may draw links between *Htt* and neuropsychiatric disorders.

As demonstrated in our study, the *Htt*-HDAC axis is modulated via chromatin reorganization, yet the current drug screenings, presented in our study for SAHA and VPA, have limited capacity to achieve layer-specific, multi-targeted effects. In this context, natural products (NPs) emerge as a promising candidate for therapeutic discovery. Unlike synthetic compounds, NPs

are naturally optimized for biological systems, particularly the immune system, which is a key player in neurological disorders such as AD and MDD (Brian, 2010; Glass et al., 2010). While the application of NPs has been limited by challenges in screening, characterization, and optimization, recent technological advances in targeted genome mining and AI-powered machine learning have opened new revenues for drug development (Atanasov et al., 2021). For instance, employing machine learning to predict bioactivities and molecular targets greatly accelerates the time needed to mine potential NP candidates. Building on these advances, our transformer-based LLM model QuokkaVision could be adapted to efficiently screen NPs for layer-specific and multi-target activities centered on the *Htt*-HDAC axis. By targeting *Htt*-HDAC axis, it is made possible for us to address both mood and cognitive symptoms in effort to improve the quality of life of aging populations, which is a critical advance for aging populations with comorbid AD and MDD.

4.3 Limitations and Future Directions

While our framework reveals spatially resolved epigenetic mechanisms, we note several limitations of our study. While QuokkaVision enables pathway-level interpretations and layer-specific annotations, the lack of spatial resolution and cohort diversity may restrict its generalizability. Due to time and resource constraints, we were unable to pursue higher-resolution profiling. However, our hypothesis of spatial epigenetic adaptation—which asserts that disease trajectories are associated with layer-specific chromatin states—requires single-cell, spatially-indexed data for rigorous validation. Future studies should establish multi-omics comprehensive mapping with tools such as 10x Visium (Stahl et al., 2016), multiplexed immunofluorescence (Lin et al., 2018), MERFISH (Chen et al., 2015), Slide-seq (Rodriques et al., 2019; Stickels et al., 2021), and STARMap (Wang et al., 2018; Zeng et al., 2023), to pin-point layer-specific epigenetic alterations. Moreover, methods developed to associate epigenetic modifications with aging lack the spatial and cellular resolution needed to map where and how such changes occur in the brain (Lu et al., 2019). Thus, extending our findings with spatially resolved methods will be essential to validate our hypothesis of spatial epigenetic adaptation and to capture the heterogeneity of neurodegenerative trajectories across individuals.

Furthermore, while our *C. elegans*-based pharmacological assays provided initial functional validation of the *Htt*-HDAC axis, they are limited by organism-wide drug exposure, low resolution, and modest throughput. The lack of spatial and temporal specificity makes it difficult to isolate *Htt*-specific effects, and the use of non-layer-specific, broad-spectrum HDAC inhibitors introduces potential off-target effects. Moreover, *C. elegans* is not optimal for high-throughput, human-relevant drug screening, which limits its translatability. To overcome these limitations, we propose an approach: first, use AlphaFold3 to predict *Htt* molecular conformations to identify structurally critical sites for potential drug binding (Abramson et al., 2024); next, integrate CRISPR perturbations in human iPSC-derived neurons to validate these target sites (Jinek et al., 2012); and finally, conduct high-throughput

drug screening with spatial precision and target specificity informed by predicted conformations and validated targets. These insights will be essential for refining HDACs-targeted therapies to specific neuronal populations exhibiting *Htt*-related deficits.

Another challenge remains in isolating age-related chromatin shifts from disease-specific changes. Our framework implicitly assumes a stable epigenetic baseline across age, but HDAC-modulated pathways likely evolve with neuronal aging. While our study specifically focused on male subjects in their 90s to compare control, MCI, and AD groups, temporal plasticity may confound interpretations of disease-specific signatures. We advocate for the development of an age-dependent database that integrates transcriptomic profiles, electroencephalogram (EEG) recordings, and behavioral data. This resource would be critically important for separating aging effects from true pathological regulation and for supporting more targeted interventions for cognitive and emotional decline.

Moving forward, this study reveals a unified therapeutic opportunity: targeting layer-specific epigenetic shifts in the *Htt*-HDAC network to address both cognitive and affective decline. Such an approach may pave the way for earlier interventions and help preserve the quality of life as populations age.

5. Conclusion and Perspectives

- We identified the *Htt*-HDAC network as a shared mechanism connecting AD and MDD.
- We revealed consistent dysregulation of histone-related genes across neuronal subtypes in MDD and selectively in AD.
- QuokkaVision, our transformer-based LLM model, highlighted chromatin regulation as a shared pathological mechanism.
- We positioned *Htt* as a key epigenetic hub, connecting chromatin dysregulation to synaptic and neurodevelopmental deficits; Genes downstream of *Htt* (*APBB2*, *PNISR*, and *ELAVL3*) are highly relevant to both AD and MDD.
- Using QuokkaVision, we mapped this network to layer 5 excitatory and inhibitory neurons.
- We validated *Htt* in vivo by knocking down its ortholog in AD *C. elegans* models, which delayed paralysis, extended lifespan, and reduced Aβ accumulation.
- We pharmacologically targeted the *Htt*-HDAC network using HDAC inhibitors, which improved *C. elegans* health and cognitive function.
- We emphasized the need for further subtype-specific and comprehensive multi-omics studies to fully resolve how epigenetic alterations contribute to disease trajectories.

Conflicts of Interest

The author declares no conflicts of interest.

Data Availability Statement

All datasets used in this study are obtained from public data repositories. MDD datasets are available from GEO "GSE208438" (MDD; scRNA-seq), "GSE148822" (AD; scRNA-seq), "PMC3837431" (GWAS), and "PMC6522363" (GWAS). AD datasets are available from GEO "GSE157827" (snRNA-seq) and Allen Brain Atlas "SEA-AD" (https://registry.opendata.aws/allen-sea-ad-atlas) (snATAC-seq). The pathway mask matrix used in this work is based on knowledge datasets from GSEA (http://www.gsea-msigdb.org/gsea/downloads.jsp) and Allen Brain Atlas (https://portal.brain-map.org). The UCSC Genome Browser used is from https://genome.ucsc.edu. Protein level analyses are conducted with STRING (https://string-db.org). *Htt*'s downstream network were analyzed with QIAGEN IPA (QIAGEN Inc., https://digitalinsights.qiagen.com/IPA). All other relevant data supporting the key findings of this study are available within the article.

Supplemental Data

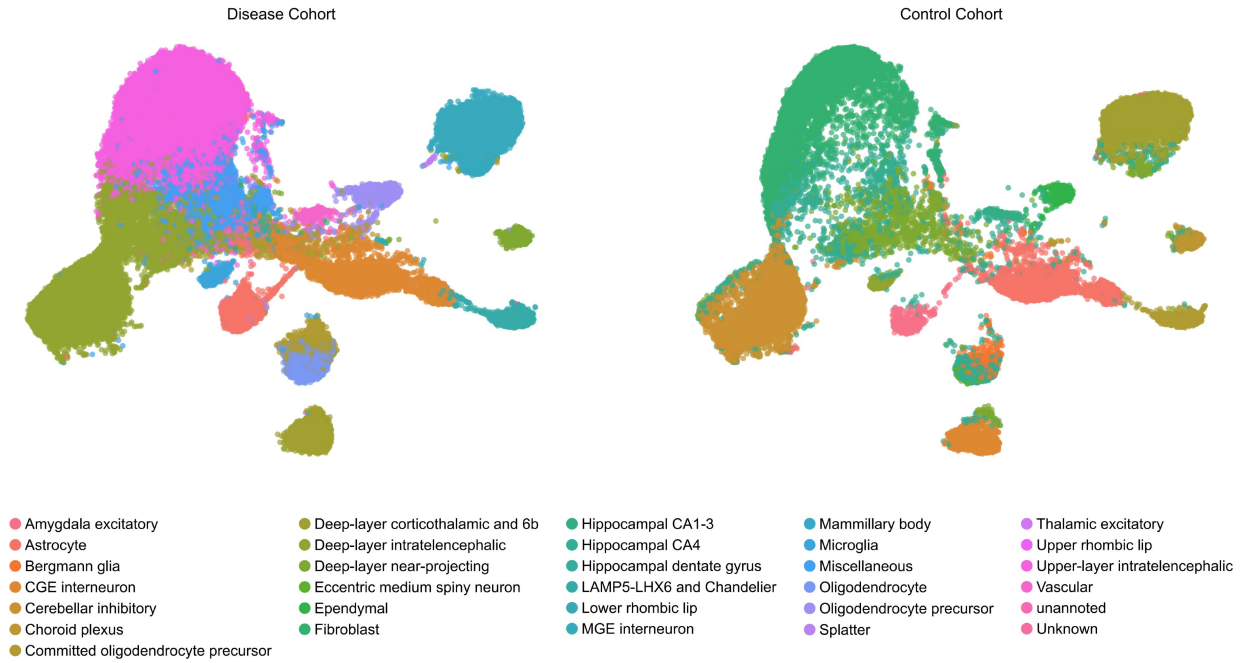


Figure S1. UMAP of MDD patients with QuokkaVision. Cells are color-coded according to predicted cell type, with two cohorts present: MDD and controls. This classification framework is used consistently in subsequent analyses.

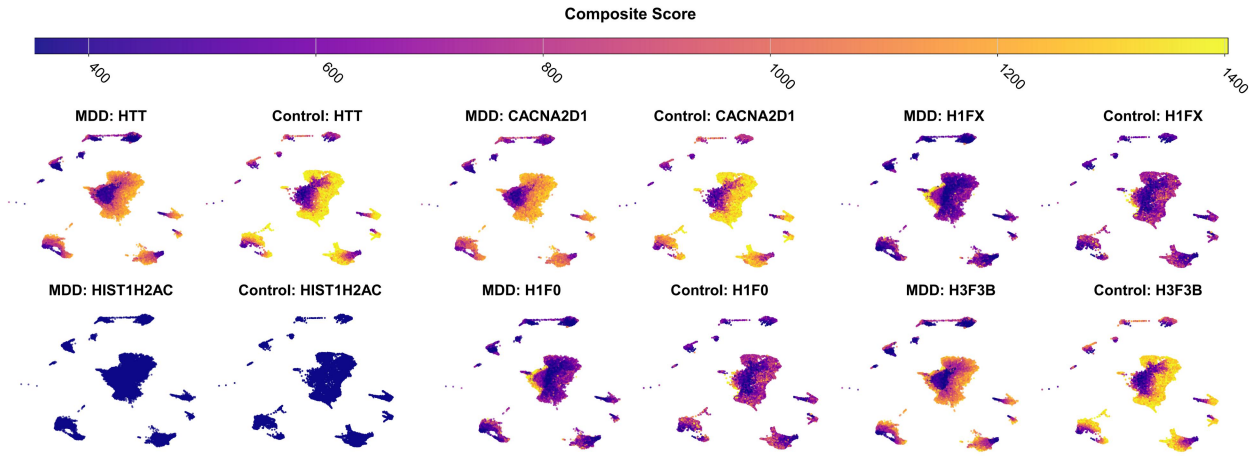


Figure S2. UMAP of *Htt*-HDAC Network Genes with QuokkaVision. Genes are mapped to UMAPs of diseased and control patients colored by composite scores, which represent the contribution of each gene across model-learned pathway activations. These scores are computed by weighing each cell's pathway activity by the gene's associated attention profile in the model.

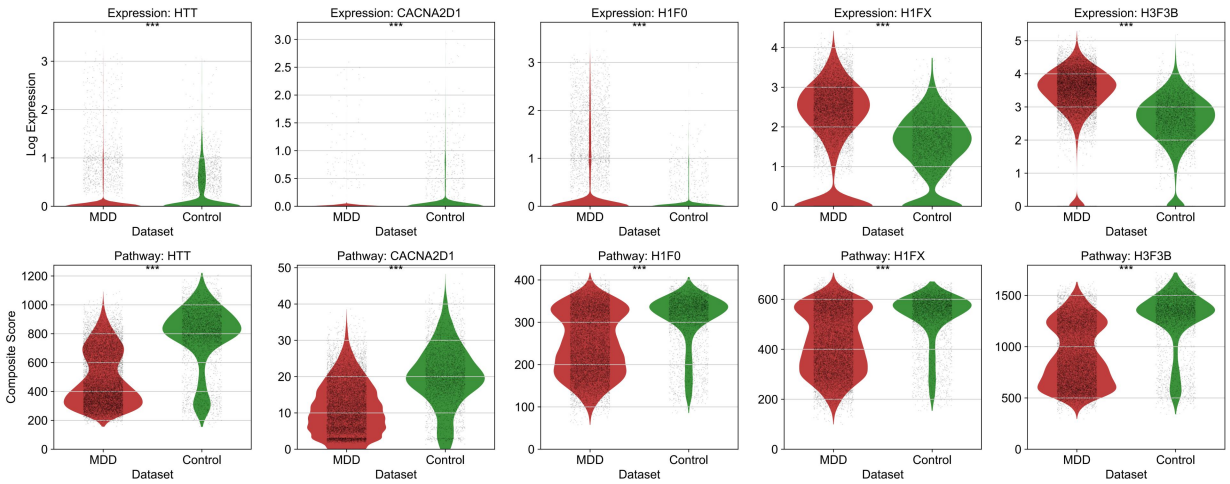


Figure S3. Violin Plots of *Htt*-HDAC Network Gene Expression. Plots show (a) gene scores recalculated using a pathway-informed masking approach (composite scores), and (b) baseline distribution of gene expression levels without pathway adjustment. Asterisks indicate statistical significance: $p < 0.001$ (***).

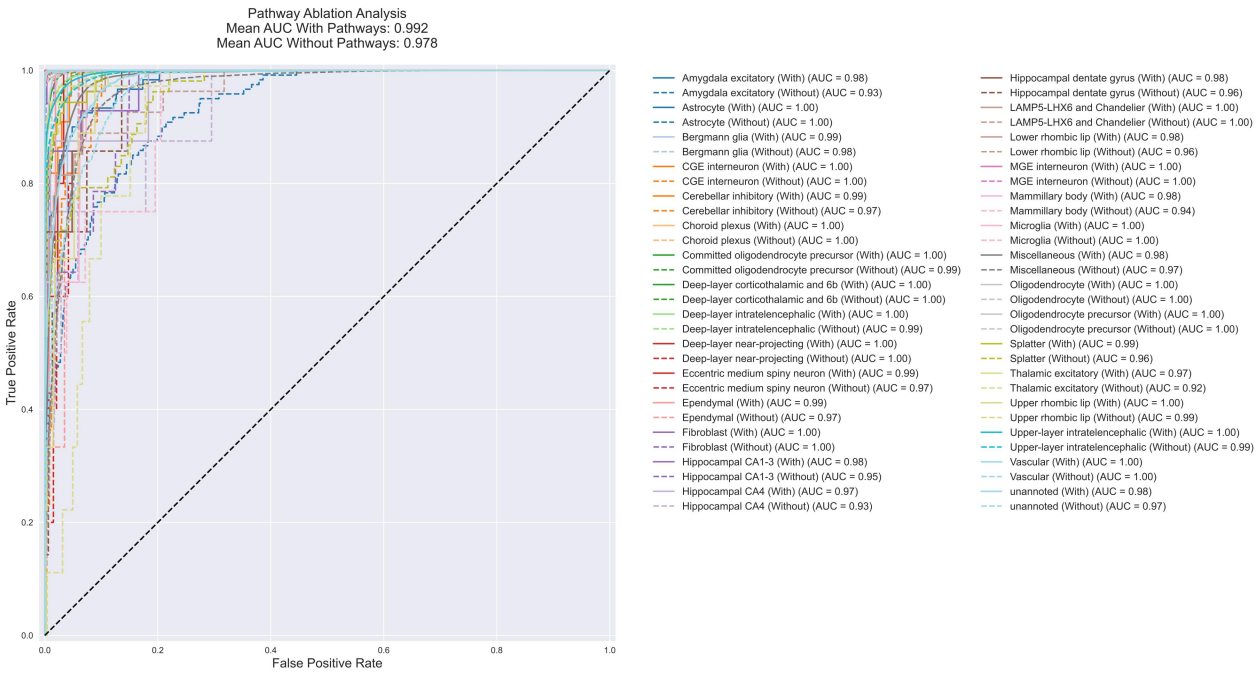


Figure S4. Validation of QuokkaVision Accuracy with and without Pathway Mask. Mean AUC with pathway mask was 0.992, compared with 0.978 without pathways (ΔAUC = 0.014, p-value: 2.32e-04).

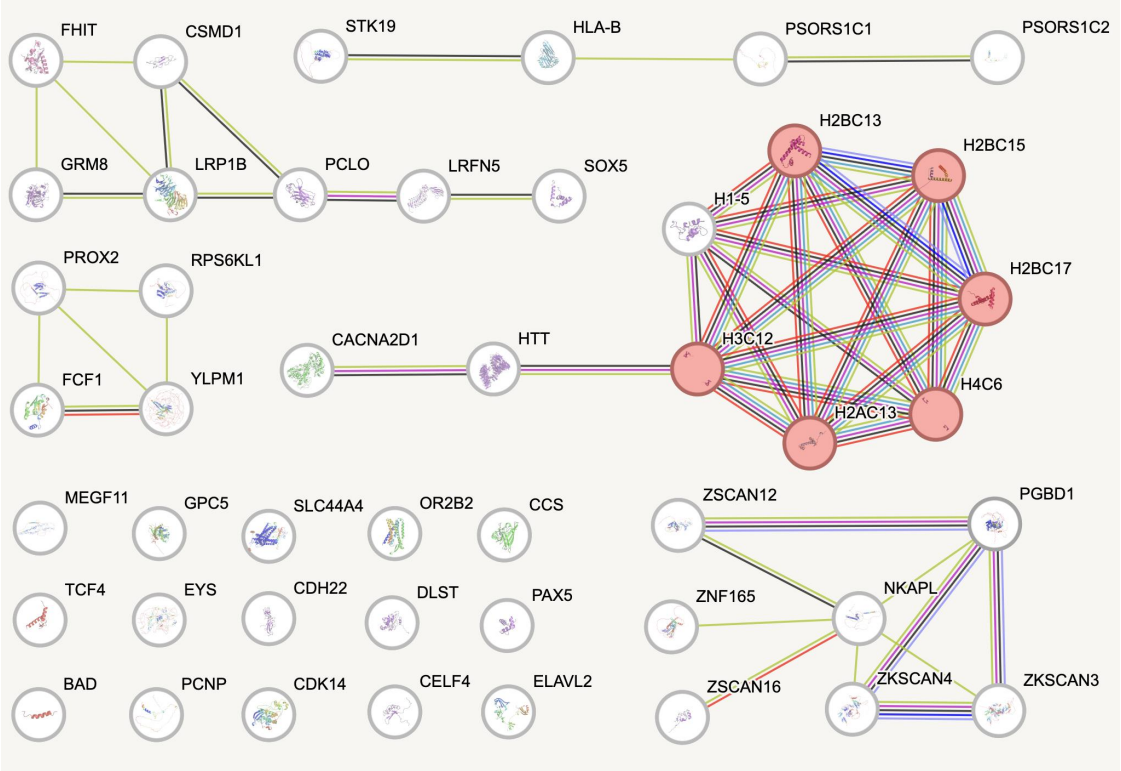


Figure S5. Protein-Protein Interaction (PPI) Network of Overlapping DEGs from Two MDD GWAS Datasets. Genes are color-coded by pathways in which HDACs deacetylate histones.

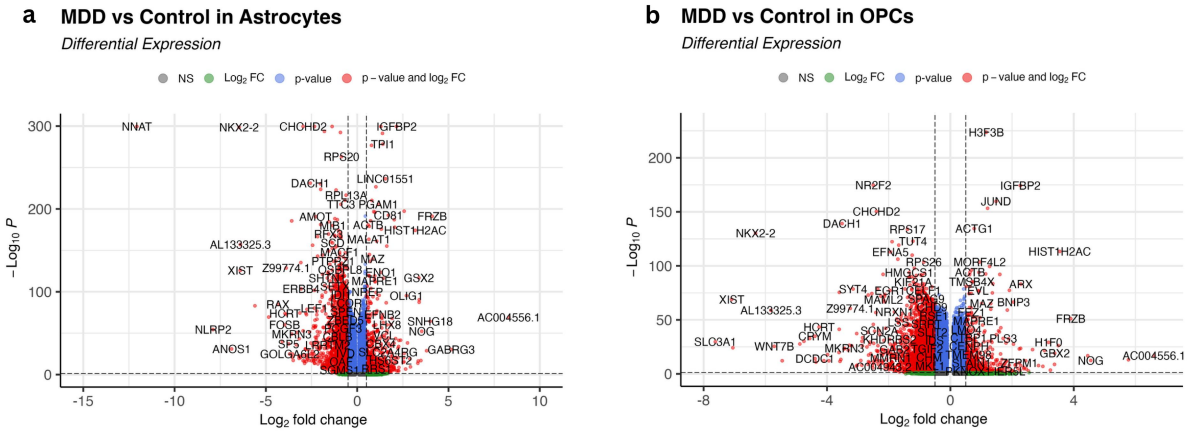


Figure S6. Transcriptomic Analysis of DEGs in MDD and Control Groups. The specific neuronal subtypes analyzed include (a) astrocytes and (b) OPCs. Genes were considered significantly differentially expressed with a p-value < 0.05 and $|\log_2 FC| \geq 0.5$.

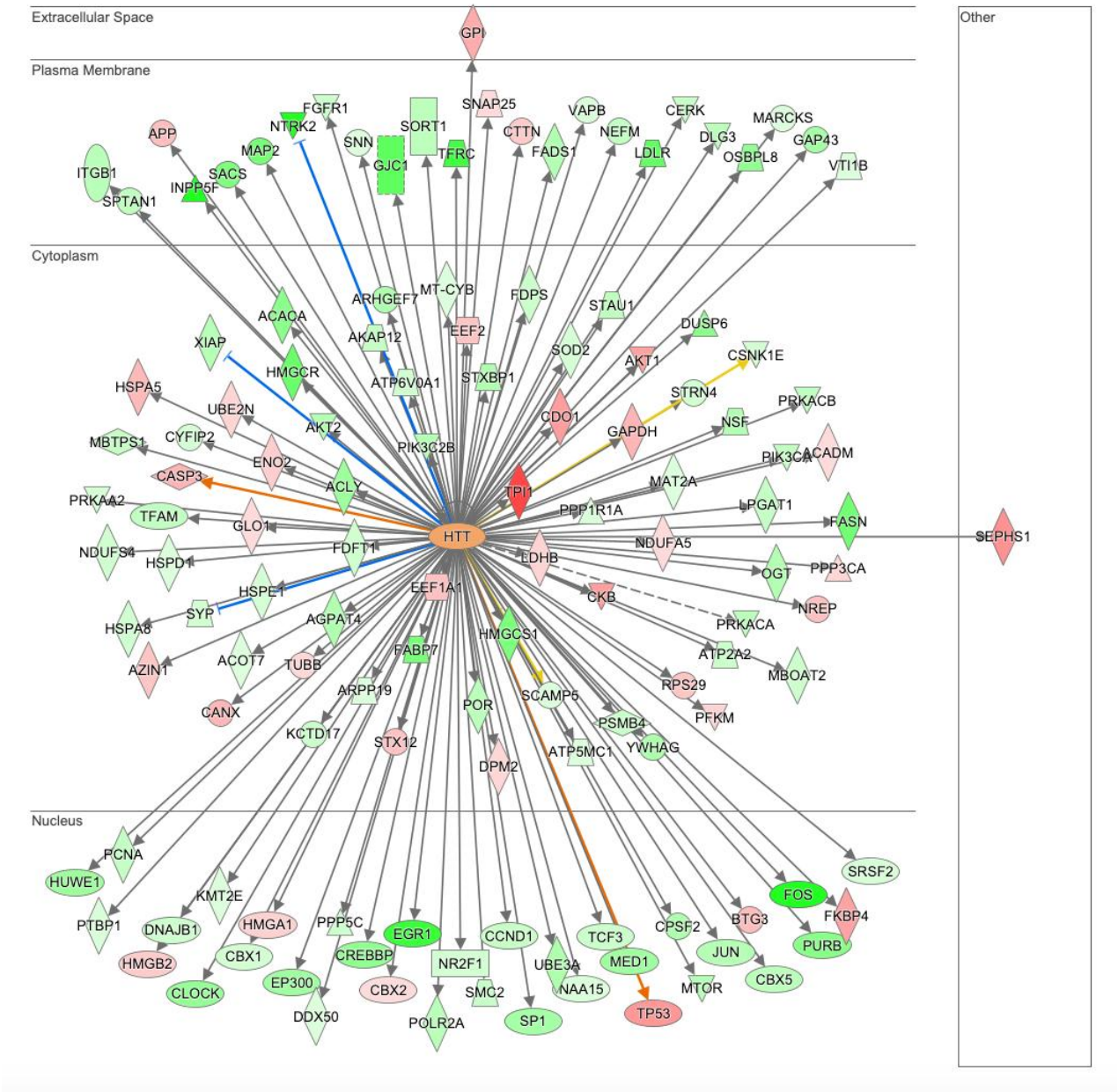


Figure S7. *Htt* Downstream Signaling Network Across Cellular Compartments. The network is mapped across extracellular space, plasma membrane, cytoplasm, and nucleus.

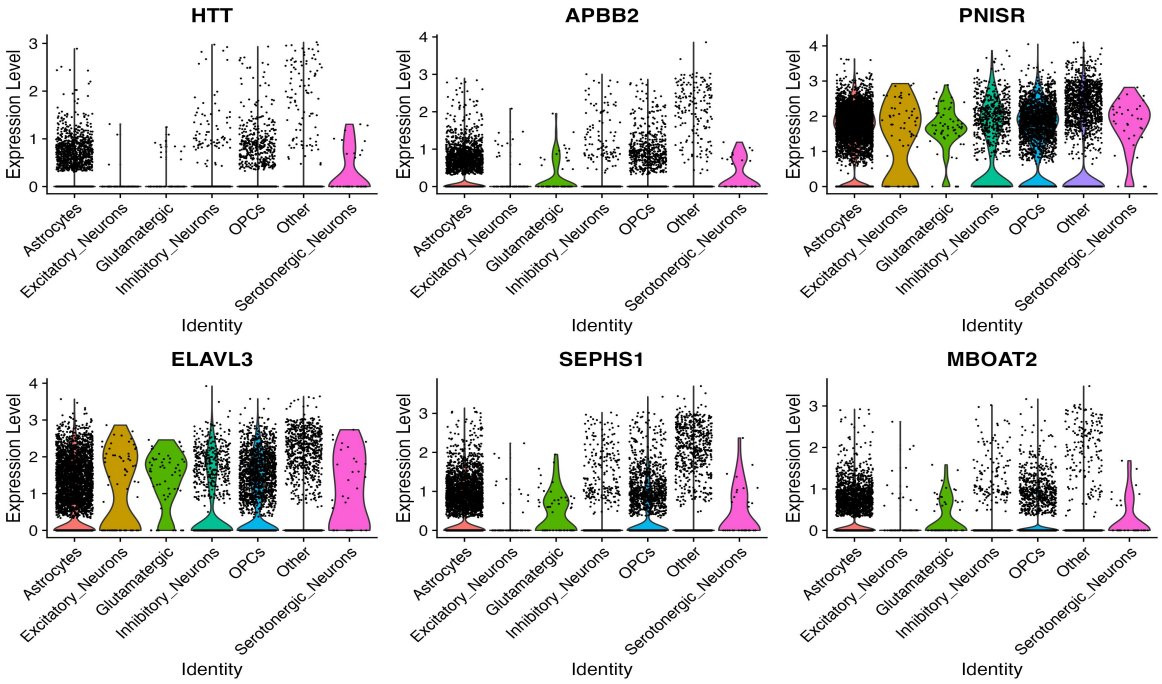


Figure S8. Expression of *Htt* Network Downstream Genes in MDD. Expression levels are shown across six different cell types: astrocytes, excitatory neurons, glutamatergic neurons, inhibitory neurons, oligodendrocyte precursor cells (OPCs), and serotonergic neurons.

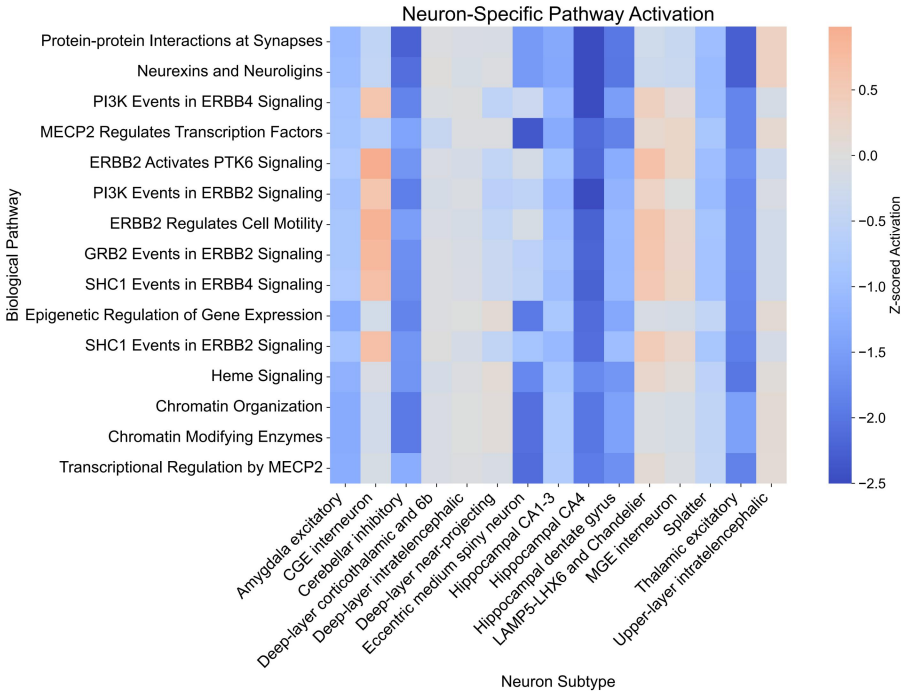


Figure S9. Neuron-specific Reactome Pathway Activation in MDD Patients. Pathway activation is mapped against a Reactome reference dataset; generated by QuokkaVision.

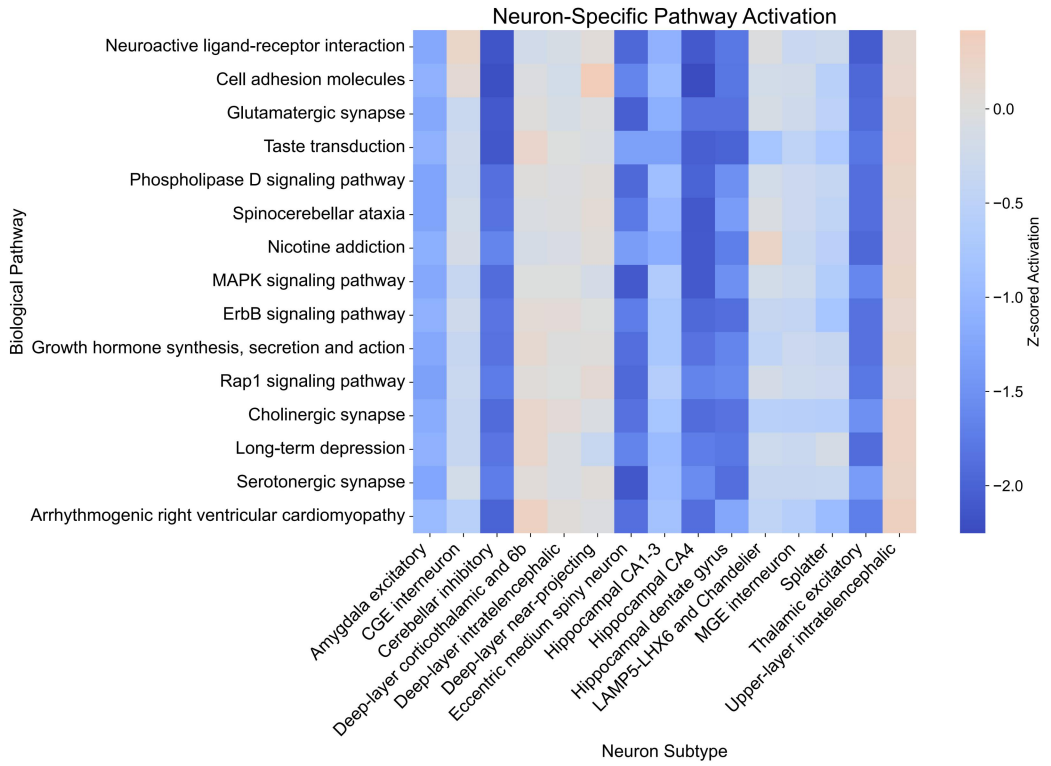


Figure S10. Neuron-specific KEGG Pathway Activation in MDD patients. Pathway activation is mapped against a KEGG pathway reference dataset; generated by QuokkaVision.

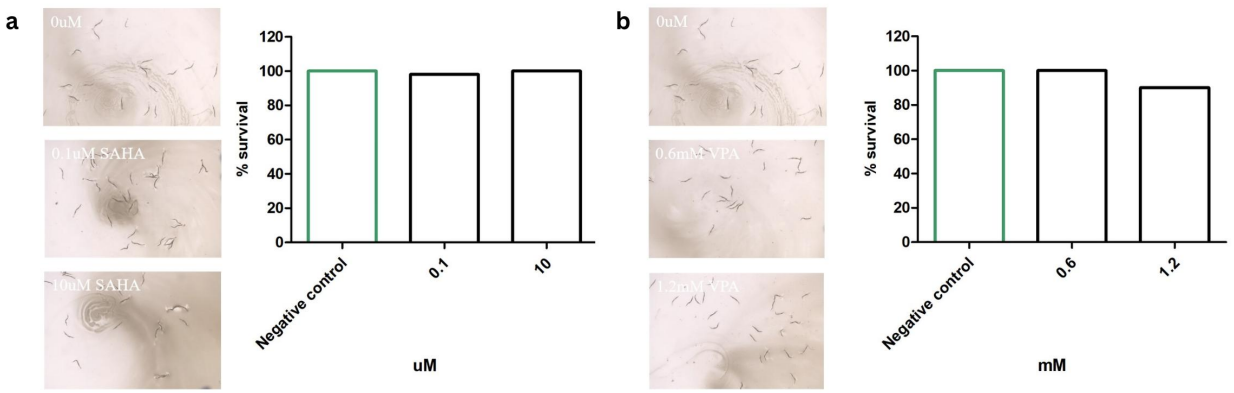


Figure S11. Dose-dependent Toxicity Assay of HDAC inhibitors in Aβ-expressing CL2006 *C. elegans* (n = 50 worms per condition). (a) SAHA: Stereomicroscopy images (left) and survival quantification (right) for vehicle and two concentrations (0.1 µM and 10 µM). (b) VPA: Stereomicroscopy images (left) and survival quantification (right) for vehicle and two concentrations (0.6 mM VPA and 1.2 mM VPA), using the same control cohort.

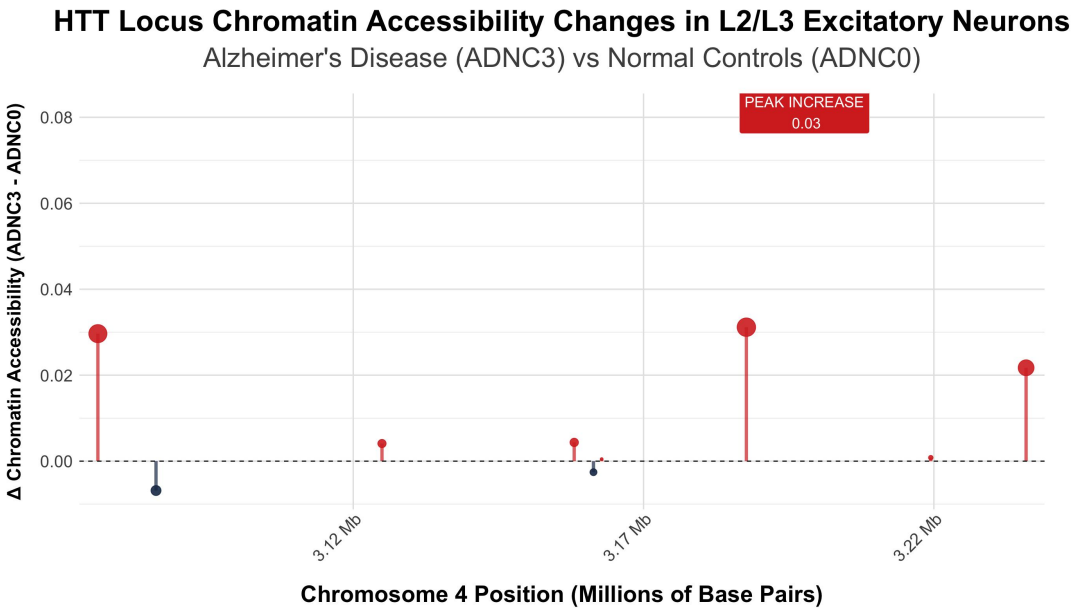


Figure S12. Chromatin Accessibility Changes at the *Htt* locus in AD patients (ADNC3 vs. Control). Changes are mapped according to chromosomal position in L2/L3 excitatory neurons.

References

Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A. J., Bambrick, J., Bodenstein, S. W., Evans, D. A., Hung, C. C., O'Neill, M., Reiman, D., Tunyasuvunakool, K., Wu, Z., Zemgulyte, A., Arvaniti, E.,…Jumper, J. M. (2024). Addendum: Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, *636*(8042), E4. https://doi.org/10.1038/s41586-024-08416-7

Adema, K., Schon, M. A., Nodine, M. D., & Kohlen, W. (2024). Lost in space: what single-cell RNA sequencing cannot tell you. *Trends Plant Sci*, *29*(9), 1018-1028. https://doi.org/10.1016/j.tplants.2024.03.010

Aguera-Ortiz, L., Garcia-Ramos, R., Grandas Perez, F. J., Lopez-Alvarez, J., Montes Rodriguez, J. M., Olazaran Rodriguez, F. J., Olivera Pueyo, J., Pelegrin Valero, C., & Porta-Etessam, J. (2021). Depression in Alzheimer's Disease: A Delphi Consensus on Etiology, Risk Factors, and Clinical Management. *Front Psychiatry*, *12*, 638651. https://doi.org/10.3389/fpsyt.2021.638651

Asmer, M. S., Kirkham, J., Newton, H., Ismail, Z., Elbayoumi, H., Leung, R. H., & Seitz, D. P. (2018). Meta-Analysis of the Prevalence of Major Depressive Disorder Among Older Adults With Dementia. *J Clin Psychiatry*, *79*(5). https://doi.org/10.4088/JCP.17r11772

Atanasov, A. G., Zotchev, S. B., Dirsch, V. M., Orhan, I. E., Banach, M., Rollinger, J. M., Barreca, D., Weckwerth, W., Bauer, R., Bayer, E. A., Majeed, M., Bishayee, A., Bochkov, V., Bonn, G. K., Braidy, N., Bucar, F., Cifuentes, A., D'Onofrio, G., Bodkin, M.,…the International Natural Product Sciences, T. (2021). Natural products in drug discovery: advances and opportunities. *Nature Reviews Drug Discovery*, *20*(3), 200-216. https://doi.org/10.1038/s41573-020-00114-z

Babulal, G. M., Roe, C. M., Stout, S. H., Rajasekar, G., Wisch, J. K., Benzinger, T. L. S., Morris, J. C., & Ances, B. M. (2020). Depression is Associated with Tau and Not Amyloid Positron Emission Tomography in Cognitively Normal Adults. *J Alzheimers Dis*, *74*(4), 1045-1055. https://doi.org/10.3233/JAD-191078

Bains, N., & Abdijadid, S. (2025). Major Depressive Disorder. In *StatPearls*. https://www.ncbi.nlm.nih.gov/pubmed/32644504

Barnes, D. E., Yaffe, K., Byers, A. L., McCormick, M., Schaefer, C., & Whitmer, R. A. (2012). Midlife vs late-life depressive symptoms and risk of dementia: differential effects for Alzheimer disease and vascular dementia. *Arch Gen Psychiatry*, *69*(5), 493-498. https://doi.org/10.1001/archgenpsychiatry.2011.1481

Bartels, C., Wagner, M., Wolfsgruber, S., Ehrenreich, H., Schneider, A., & Alzheimer's Disease Neuroimaging, I. (2018). Impact of SSRI Therapy on Risk of Conversion From Mild Cognitive Impairment to Alzheimer's Dementia in Individuals With Previous Depression. *Am J Psychiatry*, *175*(3), 232-241. https://doi.org/10.1176/appi.ajp.2017.17040404

Blier, P., & El Mansari, M. (2013). Serotonin and beyond: therapeutics for major depression. *Philos Trans R Soc Lond B Biol Sci*, *368*(1615), 20120536. https://doi.org/10.1098/rstb.2012.0536

Botto, R., Callai, N., Cermelli, A., Causarano, L., & Rainero, I. (2022). Anxiety and depression in Alzheimer's disease: a systematic review of pathogenetic mechanisms and relation to cognitive decline. *Neurol Sci*, *43*(7), 4107-4124. https://doi.org/10.1007/s10072-022-06068-x

Brauer, R., Lau, W. C. Y., Hayes, J. F., Man, K. K. C., Osborn, D. P. J., Howard, R., Kim, J., & Wong, I. C. K. (2019). Trazodone use and risk of dementia: A population-based cohort study. *PLoS Med*, *16*(2), e1002728. https://doi.org/10.1371/journal.pmed.1002728

Brian, E. L. (2010). The Concept of Depression as a Dysfunction of the Immune System. *Current Immunology Reviews (Discontinued)*, *6*(3), 205-212. https://doi.org/http://dx.doi.org/10.2174/157339510791823835

Brixi, G., Durrant, M. G., Ku, J., Poli, M., Brockman, G., Chang, D., Gonzalez, G. A., King, S. H., Li, D. B., Merchant, A. T., Naghipourfar, M., Nguyen, E., Ricci-Tam, C., Romero, D. W., Sun, G., Taghibakshi, A., Vorontsov, A., Yang, B., Deng, M.,…Hie, B. L. (2025). Genome modeling and design across all domains of life with Evo 2. *bioRxiv*, 2025.2002.2018.638918. https://doi.org/10.1101/2025.02.18.638918

Chakraborty, S., Lennon, J. C., Malkaram, S. A., Zeng, Y., Fisher, D. W., & Dong, H. (2019). Serotonergic system, cognition, and BPSD in Alzheimer's disease. *Neurosci Lett*, *704*, 36-44. https://doi.org/10.1016/j.neulet.2019.03.050

Chen, G., Huang, L. D., Jiang, Y. M., & Manji, H. K. (1999). The mood-stabilizing agent valproate inhibits the activity of glycogen synthase kinase-3. *J Neurochem*, *72*(3), 1327-1330. https://doi.org/10.1046/j.1471-4159.2000.0721327.x

Chen, J., Xu, H., Tao, W., Chen, Z., Zhao, Y., & Han, J. J. (2023). Transformer for one stop interpretable cell type annotation. *Nat Commun*, *14*(1), 223. https://doi.org/10.1038/s41467-023-35923-4

Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S., & Zhuang, X. (2015). RNA imaging. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science*, *348*(6233), aaa6090. https://doi.org/10.1126/science.aaa6090

Chen, Y., Zhou, F., Lu, W., Zeng, W., Wang, X., & Xie, J. (2022). Identification of potential Mitogen-Activated Protein Kinase-related key genes and regulation networks in molecular subtypes of major depressive disorder. *Front Psychiatry*, *13*, 1004945. https://doi.org/10.3389/fpsyt.2022.1004945

Cirrito, J. R., Wallace, C. E., Yan, P., Davis, T. A., Gardiner, W. D., Doherty, B. M., King, D., Yuede, C. M., Lee, J. M., & Sheline, Y. I. (2020). Effect of escitalopram on Abeta levels and plaque load in an Alzheimer mouse model. *Neurology*, *95*(19), e2666-e2674. https://doi.org/10.1212/WNL.0000000000010733

Cui, L., Li, S., Wang, S., Wu, X., Liu, Y., Yu, W., Wang, Y., Tang, Y., Xia, M., & Li, B. (2024). Major depressive disorder: hypothesis, mechanism, prevention and treatment. *Signal Transduct Target Ther*, *9*(1), 30. https://doi.org/10.1038/s41392-024-01738-y

Dalla-Torre, H., Gonzalez, L., Mendoza-Revilla, J., Lopez Carranza, N., Grzywaczewski, A. H., Oteri, F., Dallago, C., Trop, E., de Almeida, B. P., Sirelkhatim, H., Richard, G., Skwark, M., Beguir, K., Lopez, M., & Pierrot, T. (2025). Nucleotide Transformer: building and evaluating robust foundation models for human genomics. *Nat Methods*, *22*(2), 287-297. https://doi.org/10.1038/s41592-024-02523-z

Depression Treatment and Aβ Dynamics: A Study of Alzheimer's Disease Risk (ABD Study). (2021). https://clinicaltrials.gov/study/NCT05004987

Du, Y., Du, Y., Zhang, Y., Huang, Z., Fu, M., Li, J., Pang, Y., Lei, P., Wang, Y. T., Song, W., He, G., & Dong, Z. (2019). MKP-1 reduces Abeta generation and alleviates cognitive impairments in Alzheimer's disease models. *Signal Transduct Target Ther*, *4*, 58. https://doi.org/10.1038/s41392-019-0091-4

Dudas, R., Malouf, R., McCleery, J., & Dening, T. (2018). Antidepressants for treating depression in dementia. *Cochrane Database Syst Rev*, *8*(8), CD003944. https://doi.org/10.1002/14651858.CD003944.pub2

Fong, S., Teo, E., Ng, L. F., Chen, C. B., Lakshmanan, L. N., Tsoi, S. Y., Moore, P. K., Inoue, T., Halliwell, B., & Gruber, J. (2016). Energy crisis precedes global metabolic failure in a novel Caenorhabditis elegans Alzheimer Disease model. *Sci Rep*, *6*, 33781. https://doi.org/10.1038/srep33781

Garber, A., Weingarten, L. S., Abreu, N. J., Elloumi, H. Z., Haack, T., Hildebrant, C., Martinez-Gil, N., Mathews, J., Muller, A. J., Valenzuela Palafoll, I., Steigerwald, C., & Chung, W. K. (2024). Rare predicted deleterious FEZF2 variants are associated with a neurodevelopmental phenotype. *Am J Med Genet A*, *194*(7), e63578. https://doi.org/10.1002/ajmg.a.63578

Gauthier, L. R., Charrin, B. C., Borrell-Pages, M., Dompierre, J. P., Rangone, H., Cordelieres, F. P., De Mey, J., MacDonald, M. E., Lessmann, V., Humbert, S., & Saudou, F. (2004). Huntingtin controls neurotrophic support and survival of neurons by enhancing BDNF vesicular transport along microtubules. *Cell*, *118*(1), 127-138. https://doi.org/10.1016/j.cell.2004.06.018

Gibson, J., Russ, T. C., Adams, M. J., Clarke, T. K., Howard, D. M., Hall, L. S., Fernandez-Pujals, A. M., Wigmore, E. M., Hayward, C., Davies, G., Murray, A. D., Smith, B. H., Porteous, D. J., Deary, I. J., & McIntosh, A. M. (2017). Assessing the presence of shared genetic architecture between Alzheimer's disease and major depressive disorder using genome-wide association data. *Transl Psychiatry*, *7*(4), e1094. https://doi.org/10.1038/tp.2017.49

Glass, C. K., Saijo, K., Winner, B., Marchetto, M. C., & Gage, F. H. (2010). Mechanisms Underlying Inflammation in Neurodegeneration. *Cell*, *140*(6), 918-934. https://doi.org/10.1016/j.cell.2010.02.016

Grigolon, G., Araldi, E., Erni, R., Wu, J. Y., Thomas, C., La Fortezza, M., Laube, B., Pohlmann, D., Stoffel, M., Zarse, K., Carreira, E. M., Ristow, M., & Fischer, F. (2022). Grainyhead 1 acts as a drug-inducible conserved transcriptional regulator linked to insulin signaling and lifespan. *Nat Commun*, *13*(1), 107. https://doi.org/10.1038/s41467-021-27732-4

Heilman, K. M., & Nadeau, S. E. (2022). Emotional and Neuropsychiatric Disorders Associated with Alzheimer's Disease. *Neurotherapeutics*, *19*(1), 99-116. https://doi.org/10.1007/s13311-021-01172-w

Heun, R., Papassotiropoulos, A., Jessen, F., Maier, W., & Breitner, J. C. (2001). A family study of Alzheimer disease and early- and late-onset depression in elderly patients. *Arch Gen Psychiatry*, *58*(2), 190-196. https://doi.org/10.1001/archpsyc.58.2.190

Holmquist, S., Nordstrom, A., & Nordstrom, P. (2020). The association of depression with subsequent dementia diagnosis: A Swedish nationwide cohort study from 1964 to 2016. *PLoS Med*, *17*(1), e1003016. https://doi.org/10.1371/journal.pmed.1003016

Howard, D. M., Adams, M. J., Clarke, T. K., Hafferty, J. D., Gibson, J., Shirali, M., Coleman, J. R. I., Hagenaars, S. P., Ward, J., Wigmore, E. M., Alloza, C., Shen, X., Barbu, M. C., Xu, E. Y., Whalley, H. C., Marioni, R. E., Porteous, D. J., Davies, G., Deary, I. J.,…McIntosh, A. M. (2019). Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nat Neurosci*, *22*(3), 343-352. https://doi.org/10.1038/s41593-018-0326-7

Huang, K., Xiao, C., Glass, L. M., & Sun, J. (2021). MolTrans: Molecular Interaction Transformer for drug-target interaction prediction. *Bioinformatics*, *37*(6), 830-836. https://doi.org/10.1093/bioinformatics/btaa880

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, *337*(6096), 816-821. https://doi.org/10.1126/science.1225829

Kilgore, M., Miller, C. A., Fass, D. M., Hennig, K. M., Haggarty, S. J., Sweatt, J. D., & Rumbaugh, G. (2010). Inhibitors of class 1 histone deacetylases reverse contextual memory deficits in a mouse model of Alzheimer's disease. *Neuropsychopharmacology*, *35*(4), 870-880. https://doi.org/10.1038/npp.2009.197

Kim, H., Jeong, W., Kwon, J., Kim, Y., Park, E. C., & Jang, S. I. (2021). Association between depression and the risk of Alzheimer's disease using the Korean National Health Insurance Service-Elderly Cohort. *Sci Rep*, *11*(1), 22591. https://doi.org/10.1038/s41598-021-02201-6

Knight, E. L., Giuliano, R. J., Shank, S. W., Clarke, M. M., & Almeida, D. M. (2020). Parasympathetic and sympathetic nervous systems interactively predict change in cognitive functioning in midlife adults. *Psychophysiology*, *57*(10), e13622. https://doi.org/10.1111/psyp.13622

Lau, S. F., Cao, H., Fu, A. K. Y., & Ip, N. Y. (2020). Single-nucleus transcriptome analysis reveals dysregulation of angiogenic endothelial cells and neuroprotective glia in Alzheimer's disease. *Proc Natl Acad Sci U S A*, *117*(41), 25800-25809. https://doi.org/10.1073/pnas.2008762117

Lin, J. R., Izar, B., Wang, S., Yapp, C., Mei, S., Shah, P. M., Santagata, S., & Sorger, P. K. (2018). Highly multiplexed immunofluorescence imaging of human tissues and tumors using t-CyCIF and conventional optical microscopes. *Elife*, *7*. https://doi.org/10.7554/eLife.31657

Liu, S., & Trapnell, C. (2016). Single-cell transcriptome sequencing: recent advances and remaining challenges. *F1000Res*, *5*. https://doi.org/10.12688/f1000research.7223.1

Lu, A. T., Quach, A., Wilson, J. G., Reiner, A. P., Aviv, A., Raj, K., Hou, L., Baccarelli, A. A., Li, Y., Stewart, J. D., Whitsel, E. A., Assimes, T. L., Ferrucci, L., & Horvath, S. (2019). DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging (Albany NY)*, *11*(2), 303-327. https://doi.org/10.18632/aging.101684

Lu, K., Hong, Y., Tao, M., Shen, L., Zheng, Z., Fang, K., Yuan, F., Xu, M., Wang, C., Zhu, D., Guo, X., & Liu, Y. (2023). Depressive patient-derived GABA interneurons reveal abnormal neural activity associated with HTR2C. *EMBO Mol Med*, *15*(1), e16364. https://doi.org/10.15252/emmm.202216364

Lublin, A. L., & Link, C. D. (2013). Alzheimer's disease drug discovery: in vivo screening using Caenorhabditis elegans as a model for beta-amyloid peptide-induced toxicity. *Drug Discov Today Technol*, *10*(1), e115-e119. https://doi.org/10.1016/j.ddtec.2012.02.002

Lyketsos, C. G., Lopez, O., Jones, B., Fitzpatrick, A. L., Breitner, J., & DeKosky, S. (2002). Prevalence of neuropsychiatric symptoms in dementia and mild cognitive impairment: results from the cardiovascular health study. *JAMA*, *288*(12), 1475-1483. https://doi.org/10.1001/jama.288.12.1475

Major Depressive Disorder Working Group of the Psychiatric, G. C., Ripke, S., Wray, N. R., Lewis, C. M., Hamilton, S. P., Weissman, M. M., Breen, G., Byrne, E. M., Blackwood,

D. H., Boomsma, D. I., Cichon, S., Heath, A. C., Holsboer, F., Lucae, S., Madden, P. A., Martin, N. G., McGuffin, P., Muglia, P., Noethen, M. M.,…Sullivan, P. F. (2013). A mega-analysis of genome-wide association studies for major depressive disorder. *Mol Psychiatry*, *18*(4), 497-511. https://doi.org/10.1038/mp.2012.21

Mango, D., Saidi, A., Cisale, G. Y., Feligioni, M., Corbo, M., & Nistico, R. (2019). Targeting Synaptic Plasticity in Experimental Models of Alzheimer's Disease. *Front Pharmacol*, *10*, 778. https://doi.org/10.3389/fphar.2019.00778

Mather, M. (2024). The emotion paradox in the aging body and brain. *Ann N Y Acad Sci*, *1536*(1), 13-41. https://doi.org/10.1111/nyas.15138

Meng, J., Li, Y., Camarillo, C., Yao, Y., Zhang, Y., Xu, C., & Jiang, L. (2014). The anti-tumor histone deacetylase inhibitor SAHA and the natural flavonoid curcumin exhibit synergistic neuroprotection against amyloid-beta toxicity. *PLoS One*, *9*(1), e85570. https://doi.org/10.1371/journal.pone.0085570

Milazzo, G., Mercatelli, D., Di Muzio, G., Triboli, L., De Rosa, P., Perini, G., & Giorgi, F. M. (2020). Histone Deacetylases (HDACs): Evolution, Specificity, Role in Transcriptional Complexes, and Pharmacological Actionability. *Genes (Basel)*, *11*(5). https://doi.org/10.3390/genes11050556

Monereo-Sanchez, J., Schram, M. T., Frei, O., O'Connell, K., Shadrin, A. A., Smeland, O. B., Westlye, L. T., Andreassen, O. A., Kaufmann, T., Linden, D. E. J., & van der Meer, D. (2021). Genetic Overlap Between Alzheimer's Disease and Depression Mapped Onto the Brain. *Front Neurosci*, *15*, 653130. https://doi.org/10.3389/fnins.2021.653130

Orgeta, V., Tabet, N., Nilforooshan, R., & Howard, R. (2017). Efficacy of Antidepressants for Depression in Alzheimer's Disease: Systematic Review and Meta-Analysis. *J Alzheimers Dis*, *58*(3), 725-733. https://doi.org/10.3233/JAD-161247

Pomara, N., Bruno, D., Sarreal, A. S., Hernando, R. T., Nierenberg, J., Petkova, E., Sidtis, J. J., Wisniewski, T. M., Mehta, P. D., Pratico, D., Zetterberg, H., & Blennow, K. (2012). Lower CSF amyloid beta peptides and higher F2-isoprostanes in cognitively intact elderly individuals with major depressive disorder. *Am J Psychiatry*, *169*(5), 523-530. https://doi.org/10.1176/appi.ajp.2011.11081153

Poplawski, G. H. D., Kawaguchi, R., Van Niekerk, E., Lu, P., Mehta, N., Canete, P., Lie, R., Dragatsis, I., Meves, J. M., Zheng, B., Coppola, G., & Tuszynski, M. H. (2020). Injured adult neurons regress to an embryonic transcriptional growth state. *Nature*, *581*(7806), 77-82. https://doi.org/10.1038/s41586-020-2200-5

Rodriques, S. G., Stickels, R. R., Goeva, A., Martin, C. A., Murray, E., Vanderburg, C. R., Welch, J., Chen, L. M., Chen, F., & Macosko, E. Z. (2019). Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science*, *363*(6434), 1463-1467. https://doi.org/10.1126/science.aaw1219

Sheline, Y. I., Snider, B. J., Beer, J. C., Seok, D., Fagan, A. M., Suckow, R. F., Lee, J. M., Waligorska, T., Korecka, M., Aselcioglu, I., Morris, J. C., Shaw, L. M., & Cirrito, J. R. (2020). Effect of escitalopram dose and treatment duration on CSF Abeta levels in healthy older adults: A controlled clinical trial. *Neurology*, *95*(19), e2658-e2665. https://doi.org/10.1212/WNL.0000000000010725

Siletti, K., Hodge, R., Mossi Albiach, A., Lee, K. W., Ding, S. L., Hu, L., Lonnerberg, P., Bakken, T., Casper, T., Clark, M., Dee, N., Gloe, J., Hirschstein, D., Shapovalova, N. V., Keene, C. D., Nyhus, J., Tung, H., Yanny, A. M., Arenas, E.,…Linnarsson, S.

(2023). Transcriptomic diversity of cell types across the adult human brain. *Science*, *382*(6667), eadd7046. https://doi.org/10.1126/science.add7046

Stahl, P. L., Salmen, F., Vickovic, S., Lundmark, A., Navarro, J. F., Magnusson, J., Giacomello, S., Asp, M., Westholm, J. O., Huss, M., Mollbrink, A., Linnarsson, S., Codeluppi, S., Borg, A., Ponten, F., Costea, P. I., Sahlen, P., Mulder, J., Bergmann, O.,…Frisen, J. (2016). Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science*, *353*(6294), 78-82. https://doi.org/10.1126/science.aaf2403

Stickels, R. R., Murray, E., Kumar, P., Li, J., Marshall, J. L., Di Bella, D. J., Arlotta, P., Macosko, E. Z., & Chen, F. (2021). Highly sensitive spatial transcriptomics at near-cellular resolution with Slide-seqV2. *Nat Biotechnol*, *39*(3), 313-319. https://doi.org/10.1038/s41587-020-0739-1

Szklarczyk, D., Kirsch, R., Koutrouli, M., Nastou, K., Mehryary, F., Hachilif, R., Gable, A. L., Fang, T., Doncheva, N. T., Pyysalo, S., Bork, P., Jensen, L. J., & von Mering, C. (2023). The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res*, *51*(D1), D638-D646. https://doi.org/10.1093/nar/gkac1000

Transcranial Magnetic Stimulation for MCI: A Phase II Dose-Response Study. (2023). https://clinicaltrials.gov/study/NCT05992831

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention Is All You Need. arXiv:1706.03762. Retrieved June 01, 2017, from https://ui.adsabs.harvard.edu/abs/2017arXiv170603762V

Vorontsov, E., Bozkurt, A., Casson, A., Shaikovski, G., Zelechowski, M., Severson, K., Zimmermann, E., Hall, J., Tenenholtz, N., Fusi, N., Yang, E., Mathieu, P., van Eck, A., Lee, D., Viret, J., Robert, E., Wang, Y. K., Kunz, J. D., Lee, M. C. H.,…Fuchs, T. J. (2024). A foundation model for clinical-grade computational pathology and rare cancers detection. *Nat Med*, *30*(10), 2924-2935. https://doi.org/10.1038/s41591-024-03141-0

Wang, X., Allen, W. E., Wright, M. A., Sylwestrak, E. L., Samusik, N., Vesuna, S., Evans, K., Liu, C., Ramakrishnan, C., Liu, J., Nolan, G. P., Bava, F. A., & Deisseroth, K. (2018). Three-dimensional intact-tissue sequencing of single-cell transcriptional states. *Science*, *361*(6400). https://doi.org/10.1126/science.aat5691

Wingo, T. S., Gerasimov, E. S., Canon, S. M., Lah, J. J., Levey, A. I., & Wingo, A. P. (2023). Alzheimer's disease genetic burden is associated with mid-life depression among persons with normal cognition. *Alzheimers Dement*, *19*(3), 868-874. https://doi.org/10.1002/alz.12716

World Alzheimer Report 2024: Global changes in attitudes to dementia. (2024). [Report]. A. s. D. International.

Zeng, H., Huang, J., Zhou, H., Meilandt, W. J., Dejanovic, B., Zhou, Y., Bohlen, C. J., Lee, S. H., Ren, J., Liu, A., Tang, Z., Sheng, H., Liu, J., Sheng, M., & Wang, X. (2023). Integrative in situ mapping of single-cell transcriptional states and tissue histopathology in a mouse model of Alzheimer's disease. *Nat Neurosci*, *26*(3), 430-446. https://doi.org/10.1038/s41593-022-01251-x

Zhou, B., Zhu, Z., Ransom, B. R., & Tong, X. (2021). Oligodendrocyte lineage cells and depression. *Mol Psychiatry*, *26*(1), 103-117. https://doi.org/10.1038/s41380-020-00930-0