UC Irvine

UC Irvine Electronic Theses and Dissertations

Title

Cellular Excitability Changes in the Entorhinal Cortex and the Retrosplenial Cortex Across Stages of Pathology in Alzheimer's Disease

Permalink

https://escholarship.org/uc/item/13s3k3xr

Author

Bartas, Katrina

Publication Date

2024

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA,
IRVINE

Cellular Excitability Changes in the Entorhinal Cortex and the Retrosplenial Cortex Across Stages of Pathology in Alzheimer's Disease

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Mathematical, Computational and Systems Biology

by

Katrina Lee Bartas

Dissertation Committee:
Associate Professor Kevin T. Beier, Chair
Associate Professor Vivek Swarup
Professor Elizabeth Head

2024

Chapter 2 © 2024 eLife Sciences Publications Ltd.
All other materials © 2024 Katrina Lee Bartas

TABLE OF CONTENTS

LIST OF FIGURES iv
LIST OF TABLES v
ACKNOWLEDGEMENTS vi
VITA vii
ABSTRACT OF THE DISSERTATION viii
CHAPTER 1: Introduction and background 1
1.1 Alzheimer's disease 1
1.2 5xFAD mouse model of Alzheimer's disease 8
1.3 Viral tracing of neural circuits in the mouse brain 11
1.4 References 14
CHAPTER 2: Exploring relationships between cellular excitability and connectivity 18
2.1 Introduction 19
2.2 Methods 20
2.3 Results 31
2.4 Discussion 58
2.5 Supplementary figures 64
2.6 References 77
CHAPTER 3: Differences in cellular communication in the entorhinal cortex in AD and in 5xFAD mice 81
3.1 Introduction 81
3.2 Methods 83
3.3 Results 90
3.4 Discussion 115
3.5 Supplementary figures 120
3.6 References 130
CHAPTER 4: A comparative analysis of the retrosplenial cortex and the entorhinal cortex in AD and in 5xFAD mice 137
4.1 Introduction 137
4.2 Methods 138
4.3 Results 147
4.4 Discussion 165
4.5 Supplementary figures 167
4.6 References 179
CHAPTER 5: Conclusion and future directions 181

5.1 Discussion 181
5.2 Future directions 182
5.3 References 186

LIST OF FIGURES

		Page
Figure 2.1	RABV input tracing to VTA	35
Figure 2.2	PCA to deconvolve variation	38
Figure 2.3	Dimensionality reduction analysis: drugs	44
Figure 2.4	Dimensionality reduction analysis: ketamine	48
Figure 2.5	GO analysis of top 50 correlated genes	52
Figure 2.6	RABV labeling correlated to gene classes	55
Figure 2.7	Inhibition of *Cacna1e*	58
Figure 3.1	ENT data characterization	91
Figure 3.2	Differential gene expression	94
Figure 3.3	Species-specific cellular communication	99
Figure 3.4	NeuronChat: glutamatergic signaling	108
Figure 3.5	Multi-omics approach	112
Figure 4.1	RABV-tracing and fiber photometry	148
Figure 4.2	Cell type annotation	150
Figure 4.3	Linear regression between ENT and RSC results	153
Figure 4.4	GO analysis in glutamatergic neurons	156
Figure 4.5	Cellular communication across time, disease progression	159
Figure 4.6	Gene regulatory networks in 5xFAD mice	163

LIST OF TABLES

		Page
Table 2.1	Brain region abbreviations	27
Table 4.1	Donor metadata	140

ACKNOWLEDGEMENTS

I would like to thank my committee members, Dr. Kevin T. Beier, Dr. Vivek Swarup and Dr. Elizabeth Head, for their support, insights, and feedback on my work. Your guidance has shaped my development as a scientist and been crucial to the completion of this work. I would also like to thank Dr. Annie Qu and Dr. Devon Lawson for being part of my advancement committee and taking the time to review my work.

To all members of the Beier lab, past and present, I would like to express my thanks for all of your hard work – this dissertation would not have been possible without your help in data generation, experimental validation, and insights during data interpretation.

In addition, I am grateful for the support of everyone in the Mathematical, Computational, and Systems Biology community at UCI; my cohort provided both encouragement and inspiration, students who graduated before me were great sources of wisdom, and Dr. Emi Embler and Karen Yvonne Martin helped with every administrative or miscellaneous issue I ever encountered.

I would like to thank Dr. Qing Nie and Dr. Arthur Lander for appointing me to the NIGMS T32 MCB Training Grant, which provided years of financial support and academic guidance.

Chapter 2 of this dissertation is a reprint of the material as it appears in: Bartas, K., Hui, M., Derdeyn, P., Tian, G., Vasquez, J.J., Azouz, G., Yamamoto, C.M. and Beier, K.T. Drug-induced changes in connectivity to midbrain dopamine cells revealed by rabies monosynaptic tracing. *Elife*, 13 (2024). This work was used with permission from eLife Sciences Publications Ltd. The co-authors listed in this publication are May Hui, Pieter Derdeyn, Guilian Tian, Jose J. Vasquez, Ghalia Azouz, Cindy M. Yamamoto, and Kevin T. Beier.

This work utilized resources of the UCI Genomics Research and Technology Hub (GRT Hub) parts of which are supported by NIH grants to the Comprehensive Cancer Center (P30CA-062203) and the UCI Skin Biology Resource Based Center (P30AR075047) at the University of California, Irvine, as well as to the GRT Hub for instrumentation (1S10OD010794-01and 1S10OD021718-01). This investigation was supported by the National Institute of General Medical Sciences, National Research Service Award GM136624 from the University of California, Irvine, Center for Complex Biological Systems.

This work utilized the infrastructure for high-performance and high-throughput computing, research data storage and analysis, and scientific software tool integration built, operated, and updated by the Research Cyberinfrastructure Center (RCIC) at the University of California, Irvine (UCI). The RCIC provides cluster-based systems, application software, and scalable storage to directly support the UCI research community. https://rcic.uci.edu

VITA

Katrina Lee Bartas

2015-17	Undergraduate Researcher, Department of Molecular, Cellular, and Developmental Biology, University of California, Santa Barbara
2018	B.S. in Biological Sciences, University of California, Santa Barbara
2018-20	Research Associate I/II, Metabolic Engineering, Lygos, Inc
2022	Genomic Safe Harbor Intern, Bioinformatics, Astellas Gene Therapies
2024	Teaching Assistant, School of Biological Sciences, University of California, Irvine
2024	Ph.D. in Mathematical, Computational, and Systems Biology, University of California, Irvine

PUBLICATIONS

Bartas K, Hui M, Zhao W, Macchia D, Nie Q, Beier KT. Analysis of changes in inter-cellular communications during Alzheimer's Disease pathogenesis reveals conserved changes in glutamatergic transmission in mice and humans. doi: 10.1101/2024.04.30.591802. *In review*.

Tian G, **Bartas K**, Hui M, Chen L, Vasquez J, Azouz G, Derdeyn P, Manville R, Ho EL, Fang AS, Li Y, Tyler I, Setola V, Aoto J, Abbott GW, Beier KT. Molecular and circuit determinants in the globus pallidus mediating control of cocaine-induced behavioral plasticity. Neuron, 2024. doi: 10.1016/j.neuron.2024.07.018.

Bartas K, Hui M, Derdeyn P, Tian G, Vasquez JJ, Azouz G, Yamamoto CM, Beier KT. Drug-induced changes in connectivity to midbrain dopamine cells revealed by rabies monosynaptic tracing. eLife, 2024. doi: 10.7554/eLife.93664.1.

Hubbard E, Galinato VM, Derdeyn P, **Bartas K**, Mahler SV, Beier KT. Neural circuit basis of adolescent THC-induced potentiation of opioid responses later in life. doi: 10.1101/2024.04.23.590773. *In review*.

Tian G, Hui M, Macchia D, Derdeyn P, Rogers A, Hubbard E, Liu C, Vasquez J, Taniguchi L, **Bartas K**, Carroll S, Beier KT. An extended amygdala-midbrain circuit controlling cocaine withdrawal-induced anxiety and reinstatement. Cell Reports, 2022. doi: 10.1016/j.celrep.2022.110775.

Zhu Z, Hubbard E, Guo X, Barbosa DAN, Popal AM, Cai C, Jiang H, Zheng Z, Lin J, Gao W, Zhang J, **Bartas K**, Macchia D, Derdeyn P, Halpern CH, Mayberg HS, Beier KT, Zhu J, Wu H. A connectomic analysis of deep brain stimulation for treatment-resistant depression. Brain Stimulation, 2021. doi: 10.1016/j.brs.2021.08.010.

ABSTRACT OF THE DISSERTATION

Cellular Excitability Changes in the Entorhinal Cortex and the Retrosplenial Cortex Across Stages of Pathology in Alzheimer's Disease

by

Katrina Lee Bartas

Doctor of Philosophy in Mathematical, Computational and Systems Biology

University of California, Irvine, 2024

Associate Professor Kevin T. Beier, Chair

Alzheimer's disease (AD) is a devastating neurodegenerative disorder that is growing in prevalence and is impactful on individuals and society as a whole. Treatment options are limited, and there is no known cure; a better understanding of this disease, particularly at its early stages, is essential for the betterment of human health. This dissertation investigates how neuronal communication is disrupted throughout disease progression in AD, how these changes may be explained molecularly, and how disease-associated changes appear to occur in a brain-region specific manner.

Chapter 1 of this work reviews prior AD research, and how methods we utilize such as neuronal tracing and transcriptomic analysis have previously been applied to the research of this disease. Following this summary of previous work, we explore a rabies-based viral tracing method, the results of which can reflect cellular excitability (Chapter 2). While this component of the dissertation did not focus on AD, the methodology and insights into cellular communication, communication between brain regions, and how gene expression can influence these factors informs the rest of this work, particularly Chapter 4.

In Chapter 3, we delve into the molecular changes present in the entorhinal cortex (ENT). We utilize a mouse model of AD, and noted cellular and neuronal communication pathways that were disrupted in disease. We saw some overlap of these changes with those observed in previously published human AD data, particularly in glutamatergic signaling.

Chapter 4 synthesizes and extends the work done in earlier chapters: we first utilize the rabies-based viral tracing method characterized in Chapter 2, the results of which identified the retrospenial cortex (RSC) as an important input region to the ENT in AD mice. We followed up with multi-omic analysis of the RSC and the ENT over time in the AD mouse brain, as well as sequencing of tissue from patients with or without cognitive deficits. Glutamatergic neurons were once again seen to have gene expression changes implicating cellular excitability in disease.

Combined, these projects implicate cellular excitability in both the ENT and the RSC, but suggest that changes in the RSC may precede those in the ENT. While further research is needed and questions remain (Chapter 5), this work has yielded insights into brain-region specific molecular changes in AD, and highlights the RSC as a promising target for future study.

CHAPTER 1: Introduction and background

Chapter 1 of this dissertation aims to provide an overview of previous research relevant to the next three chapters. Alzheimer's disease's prevalence, pathology, and available treatments are described, along with an overview of relevant previous work: transcriptomic or epigenomic analysis of postmortem tissue from AD patients and from an AD mouse model. Neuronal tracing methods are then briefly reviewed, with a focus on rabies-virus-based work and previous utilization of this technique applied to mouse models of AD.

1.1 Alzheimer's disease

Prevalence and impact

Alzheimer's disease (AD) is a progressive neurological disorder characterized by memory loss, personality changes, and other behaviors associated with cognitive decline. As of 2024, it is estimated that there are 6.9 million Americans living with AD, and this figure is projected to double by 2050 as the U.S. population ages[1]. Even more common than AD, is mild cognitive impairment (MCI), which is approximately twice as common as AD[2]. The estimated rate of conversion from MCI to AD varies based on the study, but around 10% of MCI cases convert to AD within 1.5 to 10 years[3]. Even before symptoms of MCI or AD appear, there are preclinical changes occurring in the brain. These preclinical changes can include increases in the hallmarks of AD pathology, intraneuronal tau tangles and extra cellular beta-amyloid (Aβ) plaques, which have been proposed as biomarkers of preclinical AD[4]. In the early stages of AD, tau and amyloid accumulate in brain regions including the entorhinal cortex and the hippocampus, and spread as disease develops further. This accumulation continues as the preclinical state continues through to MCI, which may eventually convert to AD dementia and thus more severe pathological changes, including neuronal degeneration[5] and worsening neuroinflammation[6]. In the later stages of AD symptoms worsen,

with reduced cognitive abilities, mood changes, and physical limitations due to neurodegeneration impacting a person's independence and ability to continue their normal activities[1].

There is an apparent gender gap in AD, with women making up 4.2 million of the 6.9 million Americans with AD, and men making up 2.7 million; this difference is projected to increase over the next few decades[1,2]. This stark difference is partially explained by the increased lifespan of women compared to men, as AD risk for both men and women increases with age. Differences in incidence between men and women of the same age are unclear, with some studies reporting differences to be insignificant[7,8], an increase in incidence in men[9], or an increase in incidence in women[10,11]. Despite this uncertainty, there are clear sex and gender differences in AD stemming from both biological factors and societal norms. For example, *APOE* genotype can strongly impact a person's risk of developing AD and seems to have a stronger effect in women; a non-biological factor, lower educational attainment, is also a risk factor for AD, which has historically more greatly impacted women, as reviewed by Rocca *et al.*[12]. Regardless of the underlying cause, the outsized impact of AD on women is further increased by caregiving responsibilities that typically fall to women, with approximately two thirds of caregivers for patients with dementia being women[1].

The impact and disruption of AD extends beyond the physical toll it takes on the patient themselves, extending to affect their quality of life, families, and caregivers. As AD dementia progresses, medical attention and caregiving needs increase as independence in daily life is lost. The economic burden of AD, including medical, non-medical, caregiving, lost productivity, and other intangible costs, was estimated to be $958 billion worldwide in 2010, has only increased in recent years with the increase in AD incidence, and is projected to increase to $9.1 trillion worldwide by 2050[13]. The impact of this disease on patients, their loved ones, and society as a

whole cannot be overstated, but despite the prevalence of AD, there is no cure and treatment options are limited.

Treatment and available therapeutics

Until 2021, the only therapeutics available for AD treatment only addressed symptoms of the disease, and there had not been any new FDA approvals in almost 20 years (citation). FDA approved drugs donepezil, rivastigmine, and galantamine inhibit the acetylcholinesterase enzyme in an effort to improve cognition of patients with AD by increasing levels of acetylcholine. Until recently, acetylcholinesterase inhibitors were typically prescribed for mild to moderate AD, while in more severe cases memantine, a NMDAR antagonist, was the drug of choice[14]. While there has since been some hope and increased availability of options with the FDA approval of three monoclonal antibodies, aducanumab, lecanemab, and most recently donanemab, these treatment options are limited in impact and can result in serious side effects, as others have reviewed[15]. Their approval sparked controversy and critique, aducanumab has since been discontinued by its manufacturer, and the remaining two approaches slow disease progression but do not reverse it. New approaches, beyond the focus of current treatments, are needed. Ongoing clinical trials offer some insight into commonly targeted pathways, with potential treatment options most commonly targeting neurotransmitter receptors, inflammation and immunity, or amyloid[16].

In the development and research of treatments for AD, it is important to consider a multitude of patient populations and therapeutic targets, as recent therapies targeting amyloid have shown to be limited in effect, and gender differences, while not thoroughly discussed in recent clinical trials, are evident: men appeared to benefit more than women from donanemab[17,18]. Additionally, a majority of patients participating in these clinical trials have been white. In the donanemab trial, 91.5% of participants were white[17], while in the lecanemab trial, 80% of

participants were white[18]. Meanwhile, the population of the United States grows increasingly diverse, with a third of Americans projected to be a race other than white by 2060[19]. Given these limitations, in addition to study of therapeutic targets other than amyloid, further study of gender-specific effects, and trials including more diverse populations would be beneficial to the therapeutic landscape in AD.

Disease progression and pathogenesis

While clinical diagnosis of AD has historically been based on symptoms of cognitive impairment, neuropathological staging based on amyloid and tau pathology have been established: Thal staging, and Braak staging, respectively[20]. Thal staging quantifies the spread of Aβ, one of the hallmarks of AD and a popular target for therapeutics. Thal staging consists of five stages, starting at phase 1, corresponding to the earliest stage of amyloid pathology, and going up to phase 5, corresponding to late stages of disease and widespread Aβ deposits. In phase 1, Aβ deposits are sparse and spatially confined, typically in the neocortex. In phase 2, Aβ plaque formation is evident in the entorhinal cortex and hippocampus. Further along in phase 3, plaques spread to the diencephalon and the striatum. In the late phases of Thal staging, accumulation of Aβ spreads further throughout the brain to regions including the cerebellum and brain stem[21]. While some presence of Aβ, is normal, Aβ accumulation is considered a biomarker for early stages of AD, and often appears before symptoms of cognitive impairment are apparent. Aβ is a product of cleavage of the amyloid precursor protein (APP), and issues arise when these proteins aggregate into plaques. This aggregation is more likely to occur with higher levels of Aβ42 rather than Aβ40[22], a longer form of Aβ by just two residues. Of particular interest in relation to these plaques, is the potential role of Aβ in glutamate excitotoxicity, with previous studies demonstrating a link

between hyperactive neurons and proximity to Aβ plaques[23], and similarly linking glutamate reuptake suppression to Aβ-mediated hyperactivation[24].

Similar to Aβ, tau is a protein that is present in cognitively normal brains. In disease, tau becomes hyperphosphorylated and aggregates into neurofibrillary tangles (NFTs) within neurons. These NFTs begin to form before cognitive impairment is evident and have a consistent pattern of spread across patients. As NFTs accumulate, neurodegeneration of affected cells eventually follows[25]. While the amyloid cascade hypothesis asserts that Aβ is the driver of NFT accumulation, this hypothesis is becoming less widely accepted in favor of more complex hypotheses[26], and the relationship between tau and amyloid is unclear. Tau pathology thus has its own system of staging distinct from amyloid; there are 6 Braak stages and these are typically referred to as stages I-VI. Like Thal staging, specific brain regions tend to be impacted earlier, and as disease progresses, NFTs spread to the rest of the brain in a somewhat predictable manner. In stage I, NFTs appear in the entorhinal cortex, and in stage II, they spread further within the entorhinal cortex and into the hippocampus. Stage III is characterized by NFTs in the inferior temporal cortex, and stage IV and stage V, by NFTs in the association cortices. The final and most severe stage, stage VI, is defined by widespread NFTs in previously mentioned brain regions, and also the establishment of NFTs in the primary sensory cortex. Tau pathology has previously been linked to decreased cognitive performance, while in the same study Aβ plaques were seen to have only a moderate correlation, making NFTs a promising therapeutic target[27]. As such, tau is the fifth most common target in AD drug development, after neurotransmitter receptors, inflammation, amyloid, and synaptic plasticity[16].

Both amyloid and tau are widely studied and highly associated with AD, but their effects on specific cell types and more general molecular mechanisms underlying AD remain unclear. To

this end and with the development and increased usage of new single-cell technologies, there has been a multitude of recent studies that aim to characterize cell-type specific changes in the brains of patients with AD compared to cognitively normal patients. This expansion of high-dimensional data has enabled rapid progress in the study of the molecular underpinnings of AD, and could be channeled toward the discovery of new biomarkers and therapeutic targets.

Previous cell-type specific study of AD using transcriptomic and epigenomic data

Previous sequencing of post-mortem tissue from AD patients has largely focused on brain regions that have previously been implicated in AD, including the prefrontal cortex[28], the ENT[29,30], and the middle temporal gyrus[31]. While studies of single-cell sequencing data in AD vary in focus, some themes and consistencies across studies have emerged. Neuroinflammation, a common therapeutic target, has proven to be of particular interest, with multiple studies delving into gene expression differences in glial cells that are activated in the brain's immune response to AD. In particular, microglia[32,33] and astrocytes[34,35], have been the focus of multiple transcriptomic studies, and, along with exploration in mouse models[36], these works have led to the identification of disease-associated microglia (DAM) and astrocytes (DAA) subtypes, and the elucidation of their transcriptomic profiles compared to cells in healthy individuals.

Given the prevalence of AD in women, sex differences are another topic of interest, which have been observed across gene expression in multiple cell types, with oligodendrocytes and neurons having particularly drastic differences in one study[37], and another highlighting sex-differences in cholinergic neurons specifically[38]. Other work has expanded insights into the differential role *APOE* genotype plays in women; *LRP10* expression in AD was found to not only have sex-specific patterns in AD, but also showed a relationship to *APOE4* dosage[39].

There has recently been such an abundance of RNA sequencing data, even single-cell RNA sequencing data, that multiple attempts at creating databases have been created to make these datasets more accessible for researchers[40–42], most recently with the development of ssREAD[43], which also includes spatial transcriptomic data. While sequencing experiments will no doubt continue and the amount of available data will increase, the aggregation of previously published datasets is a great tool, as new insights can be mined through reanalysis with new bioinformatic methods. For example, this repurposing of data has been channeled towards finding novel therapeutic targets[44] or biomarkers[45] based on proteomic and transcriptomic data.

In addition to transcriptomic studies, multi-omic studies or studies that incorporate both snRNA-seq and single-nucleus assay for transposase accessible chromatin sequencing (snATAC-seq) data have been performed to assess disruptions in epigenetic gene regulation in AD[46–48], though these multi-omic studies are not as common as single modality studies. Utilization of snATAC-seq allows for insights into differences in non-coding regions of the genome, and allows for the direct analysis of AD genome-wide association study (GWAS) loci, as done previously[46]. There have been efforts to create multimodal atlases of AD data, with the Allen Institute's Seattle AD Brain Cell Atlas (SEA-AD) incorporating multi-omic data from 84 donors with varying levels of pathology[31], and thousands of sequencing experiments performed on individuals in the Religious Orders Study (ROS) or the Rush Memory and Aging Project (MAP)[49] being valuable, publicly available resources for researchers looking to study gene expression and regulation in AD.

While post-mortem sequencing analysis of brain tissue from donors has yielded a multitude of insights into AD, these studies are inherently limited by vast variability, both genetic and environmental, between patients, and only provide a single time point per person. To gain an

understanding of disease progression in brain tissue, from preclinical to late-stage AD dementia, samples from a variety of stages are necessary. At stages before diagnosis, this is especially difficult in patients, as we cannot be certain which individuals will progress or would have progressed to MCI and then to AD. With these limitations, the utilization of animal models of AD is a natural next step, as tissue samples can be taken, and corresponding analyses done at multiple points in disease in genetically identical animals. Pathology in these transgenic models occurs in a more predictable manner, so the uncertainty of whether or not disease would have occurred is eliminated. While the translational aspect of this research adds a layer of uncertainty when applying this research back to human disease, animal models open new possibilities to probe the changes that occur in AD, especially when studying early disease states or preclinical AD.

1.2 5xFAD mouse model of Alzheimer's disease

To develop mouse models of AD, many scientists have looked towards the causes of familial AD (FAD). While genetics do play a role in late onset AD (LOAD), most cases of LOAD are sporadic and thus the cause or causes are not clear. In contrast, in FAD, autosomal dominant mutations within APP and presenilins 1 (PS1) and 2 (PS2) consistently cause AD as patients age. While NFTs are also associated with AD, models based on the proteins implicated in FAD model amyloid and not tau pathology in AD, as APP cleavage results in Aβ, and PS1 plays a role in this cleavage.

The first transgenic mouse model of AD incorporated one of these FAD APP mutations, and expression of this mutant human protein was driven by the platelet-derived growth factor-β promoter. The resulting model, the PDAPP mouse[50], showed human Aβ deposits, gliosis, reduced synaptic density, and memory deficits that worsened with age. This and other mouse models have since, despite their limitations, yielded valuable insights into AD: experimentation in the mouse

brain gave insight into the role of Aβ42 as a seed for amyloid deposition, *APOE*'s role in plaque pathology, and the investigation of immunotherapeutic approaches to AD[51].

In the years since the development of the PDAPP mouse, many more mouse models of AD have been developed. One widely used model that incorporates more mutations implicated in FAD, accelerating plaque development, is the 5xFAD mouse model, which we utilize in experiments described in chapters 3 and 4 of this dissertation.

Genetic makeup of 5xFAD compared to C57BL/6 mice

In this work, we utilized the 5xFAD mouse model of AD, which utilizes mutations in the *APP* and *PSEN1* (which encodes PS1) genes known to cause FAD. Five FAD mutations are included in transgenes in the 5xFAD mouse model, hence the name. Three of these mutations affect APP: K670N/M671L, I716V, and V717I; two affect PS1: M146L and L286V[52]. The genetic background of 5xFAD mice was originally the B6.SJL mouse, but 5xFAD mice with the popular C57BL/6 lab strain as background have been created and seen widespread use. Expression of these transgenes is driven by the neuron-specific components of the *Thy1* promoter, enabling the expressed proteins to accumulate in the brain specifically.

Pathology and disruptions occur earlier in female 5xFAD mice due to the utilization of *Thy1*. This sex difference is attributed to an estrogen response element within *Thy1*, causing even higher expression of transgenes and thus amyloid pathology in female versus male 5xFAD mice[53].

Pathogenesis and behavior of 5xFAD mice

5xFAD mouse brains reflect the amyloid pathology typically found in the brains of AD patients, and their behavior over time also reflects symptoms of AD such as memory impairment. Amyloid pathology progression reflects the pattern of spread seen in humans, with Aβ plaques first accumulating in young 5xFAD mice in parts of the cortex and subiculum, spreading further

in these regions and to the hippocampus in middle age, and eventually developing in the thalamus and brain stem in aged mice[52]. NFTs were not observed. Neuroinflammation was observed as early as 2 months of age, with activated astrocytes and microgliosis being observed with the establishment of the first Aβ plaques. Spatial memory in 5xFAD mice compared to controls was evaluated using spontaneous alternation performance in the Y-maze, with impairments in this task being evident by 4-5 months of age, reflecting the memory deficits seen in AD patients. Neuronal degeneration and loss were noted at 9 months of age, evidenced by both a reduction in synaptic marker expression, and visual inspection of and neuron counting in stained brain sections[52].

Other differences noted in 5xFAD mice compared to controls include impaired long-term potentiation (LTP) at 4 months of age, decreased weight at 12 months of age, and increasing differences in gene expression with age[54]. Unsurprisingly given differential transgene expression by sex, a previous study of many aspects of 5xFAD mouse behavior and pathology noted that differences between controls and 5xFAD mice were often driven by differences in female mice rather than males, including weight gain, motor learning, number of plaques, and microglial density[54]. This study provides further evidence that the female 5xFAD mouse phenotype is impacted earlier in age than males.

Usage of 5xFAD mice in previous transcriptomic studies

Just as there has been an explosion in single-cell RNA sequencing (scRNA-seq) studies in post-mortem tissue from AD patients, there has been a similar, though not as drastic, increase in the availability of such data from 5xFAD mice (as other models are available and also widely used). In 5xFAD mice specifically, focus on glial cells involved in the immune response to AD, astrocytes and microglia, were again a focus of multiple studies[35,36,55–59]. Microglia-focused work further explored DAMs and contrasted them to microglia clusters identified as white matter

associated (WAMs), which had different patterns of gene expression and abundance suggesting a protective role for WAMs in aging, and potential conversion of WAMs to DAMs in AD[59]. Further suggesting a protective role of some microglia populations in AD, were experiments showing that microglia depletion in AD mice could be linked to bleeding and calcification in the brain[58]. Astrocytes, specifically DAAs, were observed to have some overlap in gene expression with previously profiled DAMs, and be located adjacent to Aβ plaques in the AD mouse brain[35]. There has also been recent study of oligodendrocytes in 5xFAD mice, suggesting the existence of a subpopulation of disease-associated oligodendrocytes (DAOs), which reflected the transcriptional changes seen in DAMs and DAAs[60]. Transcriptomic analysis of other mouse models and brain regions has also been widely performed, including some studies utilizing spatial transcriptomic techniques[61–64].

1.3 Viral tracing of neural circuits in the mouse brain

Another advantage to the utilization of mouse models in AD and neuroscience research more generally is the ability to clearly define anatomical connectivity between brain regions using techniques that are not possible in human subjects. Recombinant viral vectors have proven to be useful tools in the characterization of neural circuit anatomy, providing insights into the inputs to or outputs of the injection site using retrograde or anterograde tracers, respectively. Circuit tracing was first accomplished with Herpes simplex virus (HSV)[65], but has since expanded to include many other, less cytotoxic, viruses; these viral vectors can be modified for optimal characteristics such as reduced cytotoxicity, and when used in tandem with particular transgenic mouse lines, can be utilized to obtain cell-type specific information. Commonly used viruses in circuit tracing include HSV, canine adenovirus-2 (CAV2), adeno-associated virus (AAV), rabies virus (RABV), and pseudorabies virus (PRV)[66]. Typically, in anterograde tracing, a brain region of interest is

injected with a viral tracer. After enough time has passed to allow for the virus to infect other cells and for the marker gene to be expressed, imaging and quantification of the marker, typically a fluorescent protein such as GFP or mCherry, is conducted. Brain regions with a high expression of the marker gene are determined to be outputs of the injection site, or brain regions that receive projections from the injection site. Retrograde tracing is performed in a similar manner, with the injection site being the output, and brain regions with high marker gene expression being likely inputs. These techniques have enabled the mapping of neural circuitry throughout the mouse brain, with efforts to map the entire connectome, such as the Allen Mouse Brain Connectivity Atlas[67], publicly available for download and analysis.

RABV tracing overview

In chapters 2 and 4 of this dissertation, multiple tracing experiments are performed to gain insights into neuronal communication and how it changes with experience or disease. The viral tracer we elected to use is a RABV-based tracer that allows for genetic targeting of starter cells and maps exclusively monosynaptic connections[68]. Inputs are identified via expression of a fluorescent protein only if there is a direct connection to initially infected cells, and cells are only infected to begin with if they are expressing a particular gene marker, which can be induced in a genetically defined population of interest. The specificity and efficiency of this technique enable the discovery of inputs to the injection site without the ambiguity that is inherent in anterograde tracing other retrograde tracing methods.

Previous results from RABV tracing in mouse models of AD

Given that RABV tracing experiments begin with an injection into a brain region of interest, many of these studies have focused on brain regions that have known roles in the progression of AD. As the hippocampus is a brain region affected early on in AD, a previous

mouse brain connectivity study using monosynaptic RABV tracing in the APP-KI mouse model focused on inputs to hippocampal CA1 neurons[69], finding that APP-KI mice had an overall decreased connectivity strength compared to controls. Further study of inputs to excitatory neurons in a different subregion of the hippocampus in the 5xFAD mouse model resulted in similar findings; there appeared to be a decrease in overall connectivity in the AD model mice with age and compared to controls, with authors noting sex-specific differences that varied with age: female 5xFAD mice displayed higher connectivity than males at early timepoints, but lower connectivity in middle-age[70]. Two earlier studies in the APP/PS1 mouse model of AD also focused on the hippocampus, with one study focusing solely on synaptic plasticity within the hippocampus itself. In this work, bilateral stereotaxic injections were performed to label and image hippocampal CA3 inputs to CA1 neurons, in an effort to assess dendritic spine density and morphology. There was a decrease in spine density in 6 month old APP/PS1 mice compared to controls, but morphology differences were not deemed to be of consequence[71]. The second aforementioned study focused on neurons in the hippocampus expressing somatostatin and showed a similar effect on dendritic spines, with the authors noting axon loss and decreased dendritic spine density in the mouse model of AD compared to controls[72]. As we will discuss in chapter 2 of this dissertation, RABV tracing experiments may reveal information about not only anatomical connectivity, but potential changes in the activity of input cells as well. With this potential for insights into cellular activity in addition to anatomical connectivity, and current information about this type of change in AD being difficult to study in patients, further study using this method in mouse models of AD is justified, with an expansion of study into brain regions beyond the hippocampus. Based on previous study of the ENT's role in AD and our analyses in chapter 3, the ENT was a natural target for study using RABV tracing. Utilization of this technique led to discovery of the retrosplenial cortex (RSC) as

an important input to the ENT early on in AD pathology in the 5xFAD mouse model. Thus, the focus of chapter 4 is on the RSC's gene expression in 5xFAD mice compared to that of the ENT, and its potential role in the early stages of AD.

1.4 References

1. Alzheimer's Association. 2024 Alzheimer's Disease Facts and Figures. *Alzheimers Dement*. (2024).
2. Rajan, K. B. et al. Population estimate of people with clinical Alzheimer's disease and mild cognitive impairment in the United States (2020–2060). Alzheimer's and Dementia **17**, 1966–1975 (2021).
3. Bruscoli, M. & Lovestone, S. Is MCI really just early dementia? A systematic review of conversion studies. International Psychogeriatrics vol. 16 129–140 Preprint at https://doi.org/10.1017/S1041610204000092 (2004).
4. Dubois, B. et al. Preclinical Alzheimer's disease: Definition, natural history, and diagnostic criteria. Alzheimer's and Dementia vol. 12 292–323 Preprint at https://doi.org/10.1016/j.jalz.2016.02.002 (2016).
5. DeKosky, S. T. & Scheff, S. W. Synapse loss in frontal cortex biopsies in Alzheimer's disease: Correlation with cognitive severity. Ann Neurol **27**, 457–464 (1990).
6. Calsolaro, V. & Edison, P. Neuroinflammation in Alzheimer's disease: Current evidence and future directions. Alzheimer's and Dementia vol. 12 719–732 Preprint at https://doi.org/10.1016/j.jalz.2016.02.010 (2016).
7. Hebert, L. E., Scherr, P. A., McCann, J. J., Beckett, L. A. & Evans, D. A. Is the Risk of Developing Alzheimer's Disease Greater for Women than for Men? Am J Epidemiol **153**, 132–136 (2001).
8. Kawas, C., Gray, S., Brookmeyer, R., Fozard, J. & Zonderman, A. Age-specific incidence rates of Alzheimer's disease The Baltimore Longitudinal Study of Aging. Neurology **54**, 2072–2077 (2000).
9. Matthews, F. E. et al. A two decade dementia incidence comparison from the Cognitive Function and Ageing Studies i and II. Nat Commun **7**, (2016).
10. Fratiglioni, L. et al. Very Old Women at Highest Risk of Dementia and Alzheimer's Disease: Incidence data from the Kungsholmen Project, Stockholm. Neurology **48**, 132–138 (1997).
11. Letenneur, L. et al. Are sex and educational level independent predictors of dementia and Alzheimer's disease? Incidence data from the PAQUID project. J Neurol Neurosurg Psychiatry **66**, 177–183 (1999).
12. Rocca, W. A., Mielke, M. M., Vemuri, P. & Miller, V. M. Sex and gender differences in the causes of dementia: A narrative review. Maturitas vol. 79 196–201 Preprint at https://doi.org/10.1016/j.maturitas.2014.05.008 (2014).
13. Tahami Monfared, A. A., Byrnes, M. J., White, L. A. & Zhang, Q. The Humanistic and Economic Burden of Alzheimer's Disease. Neurology and Therapy vol. 11 525–551 Preprint at https://doi.org/10.1007/s40120-022-00335-x (2022).

14. Abeysinghe, A. A. D. T., Deshapriya, R. D. U. S. & Udawatte, C. Alzheimer's disease; a review of the pathophysiological basis and therapeutic interventions. Life Sciences vol. 256 Preprint at https://doi.org/10.1016/j.lfs.2020.117996 (2020).
15. Lacorte, E. et al. Safety and Efficacy of Monoclonal Antibodies for Alzheimer's Disease: A Systematic Review and Meta-Analysis of Published and Unpublished Clinical Trials. Journal of Alzheimer's Disease vol. 87 101–129 Preprint at https://doi.org/10.3233/JAD-220046 (2022).
16. Cummings, J. et al. Alzheimer's disease drug development pipeline: 2024. Alzheimer's and Dementia: Translational Research and Clinical Interventions vol. 10 Preprint at https://doi.org/10.1002/trc2.12465 (2024).
17. Sims, J. R. et al. Donanemab in Early Symptomatic Alzheimer Disease: The TRAILBLAZER-ALZ 2 Randomized Clinical Trial. JAMA **330**, 512–527 (2023).
18. van Dyck, C. H. et al. Lecanemab in Early Alzheimer's Disease. New England Journal of Medicine **388**, 9–21 (2023).
19. Vespa, J., Medina, L. & Armstrong, D. M. Demographic Turning Points for the United States: Population Projections for 2020 to 2060 Population Estimates and Projections Current Population Reports. www.census.gov/programs-surveys/popproj.
20. Therriault, J. et al. Staging of Alzheimer's disease: past, present, and future perspectives. Trends in Molecular Medicine vol. 28 726–741 Preprint at https://doi.org/10.1016/j.molmed.2022.05.008 (2022).
21. Hampel, H. et al. The Amyloid-β Pathway in Alzheimer's Disease. Mol Psychiatry **26**, 5481–5503 (2021).
22. Gu, L. & Guo, Z. Alzheimer's Aβ42 and Aβ40 peptides form interlaced amyloid fibrils. J Neurochem **126**, 305–311 (2013).
23. Busche, M. A. et al. Clusters of Hyperactive Neurons Near Amyloid Plaques in a Mouse Model of Alzheimer's Disease. Science (1979) **321**, 1686–1689 (2008).
24. Zott, B. et al. A vicious cycle of β amyloid–dependent neuronal hyperactivation. Science (1979) **365**, 559–565 (2019).
25. Iqbal, K. et al. Tau pathology in Alzheimer disease and other tauopathies. Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease **1739**, 198–210 (2005).
26. Herrup, K. The case for rejecting the amyloid cascade hypothesis. Nat Neurosci **18**, 794–799 (2015).
27. Bejanin, A. et al. Tau pathology and neurodegeneration contribute to cognitive impairment in Alzheimer's disease. Brain **140**, 3286–3300 (2017).
28. Mathys, H. et al. Single-cell atlas reveals correlates of high cognitive function, dementia, and resilience to Alzheimer's disease pathology. Cell **186**, 4365-4385.e27 (2023).
29. Grubman, A. et al. A single-cell atlas of entorhinal cortex from individuals with Alzheimer's disease reveals cell-type-specific gene expression regulation. Nat Neurosci **22**, 2087–2097 (2019).
30. Smith, A. M. et al. Diverse human astrocyte and microglial transcriptional responses to Alzheimer's pathology. Acta Neuropathol **143**, 75–91 (2022).
31. Gabitto[1], M. I. et al. Integrated multimodal cell atlas of Alzheimer's disease. doi:10.1101/2023.05.08.539485.
32. Alsema, A. M. et al. Profiling Microglia From Alzheimer's Disease Donors and Non-demented Elderly in Acute Human Postmortem Cortical Tissue. Front Mol Neurosci **13**, (2020).
33. Olah, M. et al. Single cell RNA sequencing of human microglia uncovers a subset associated with Alzheimer's disease. Nat Commun **11**, 6129 (2020).

34. Pushparaj, P. N., Kalamegam, G., Wali Sait, K. H. & Rasool, M. Decoding the Role of Astrocytes in the Entorhinal Cortex in Alzheimer's Disease Using High-Dimensional Single-Nucleus RNA Sequencing Data and Next-Generation Knowledge Discovery Methodologies: Focus on Drugs and Natural Product Remedies for Dementia. Front Pharmacol **12**, (2022).
35. Habib, N. et al. Disease-associated astrocytes in Alzheimer's disease and aging. Nat Neurosci **23**, 701–706 (2020).
36. Keren-Shaul, H. et al. A Unique Microglia Type Associated with Restricting Development of Alzheimer's Disease. Cell **169**, 1276-1290.e17 (2017).
37. Mathys, H. et al. Single-cell transcriptomic analysis of Alzheimer's disease. Nature **570**, 332–337 (2019).
38. Shulman, D. et al. Sex-specific declines in cholinergic-targeting tRNA fragments in the nucleus accumbens in Alzheimer's disease. Alzheimer's & Dementia **19**, 5159–5172 (2023).
39. Guo, L. et al. Sex specific molecular networks and key drivers of Alzheimer's disease. Mol Neurodegener **18**, 39 (2023).
40. Li, X. W. et al. SCAD-Brain: a public database of single cell RNA-seq data in human and mouse brains with Alzheimer's disease. Front Aging Neurosci **15**, (2023).
41. Zhou, Y. et al. The Alzheimer's Cell Atlas (TACA): A single-cell molecular map for translational therapeutics accelerator in Alzheimer's disease. Alzheimer's and Dementia: Translational Research and Clinical Interventions **8**, (2022).
42. Jiang, J., Wang, C., Qi, R., Fu, H. & Ma, Q. scREAD: A Single-Cell RNA-Seq Database for Alzheimer's Disease. iScience **23**, (2020).
43. Wang, C. et al. A single-cell and spatial RNA-seq database for Alzheimer's disease (ssREAD). Nat Commun **15**, (2024).
44. Ou, Y.-N. et al. Identification of novel drug targets for Alzheimer's disease by integrating genetics and proteomes from brain and blood. Mol Psychiatry **26**, 6065–6073 (2021).
45. Yu, Q.-S., Feng, W.-Q., Shi, L.-L., Niu, R.-Z. & Liu, J. Integrated Analysis of Cortex Single-Cell Transcriptome and Serum Proteome Reveals the Novel Biomarkers in Alzheimer's Disease. Brain Sci **12**, 1022 (2022).
46. Morabito, S. et al. Single-nucleus chromatin accessibility and transcriptomic characterization of Alzheimer's disease. Nat Genet **53**, 1143–1155 (2021).
47. Gamache, J. et al. Integrative single-nucleus multi-omics analysis prioritizes candidate cis and trans regulatory networks and their target genes in Alzheimer's disease brains. Cell Biosci **13**, 185 (2023).
48. Anderson, A. G. et al. Single nucleus multiomics identifies ZEB1 and MAFB as candidate regulators of Alzheimer's disease-specific cis-regulatory elements. Cell Genomics **3**, 100263 (2023).
49. De Jager, P. L. et al. A multi-omic atlas of the human frontal cortex for aging and Alzheimer's disease research. Sci Data **5**, 180142 (2018).
50. Games, D. et al. Alzheimer-type neuropathology in transgenic mice overexpressing V717F β-amyloid precursor protein. Nature **373**, 523–527 (1995).
51. Elder, G. A., Gama Sosa, M. A. & De Gasperi, R. Transgenic Mouse Models of Alzheimer's Disease. Mount Sinai Journal of Medicine: A Journal of Translational and Personalized Medicine **77**, 69–81 (2010).
52. Oakley, H. et al. Intraneuronal β-amyloid aggregates, neurodegeneration, and neuron loss in transgenic mice with five familial Alzheimer's disease mutations: Potential factors in amyloid plaque formation. Journal of Neuroscience **26**, 10129–10140 (2006).

53. Sadleir, K. R., Eimer, W. A., Cole, S. L. & Vassar, R. Aβ reduction in BACE1 heterozygous null 5XFAD mice is associated with transgenic APP level. Mol Neurodegener **10**, 1 (2015).
54. Forner, S. et al. Systematic phenotyping and characterization of the 5xFAD mouse model of Alzheimer's disease. Sci Data **8**, (2021).
55. Wang, H. Microglia Heterogeneity in Alzheimer's Disease: Insights From Single-Cell Technologies. Front Synaptic Neurosci **13**, (2021).
56. Fatmi, M. K. et al. Single-Cell RNA-seq reveals transcriptomic modulation of Alzheimer's disease by activated protein C. Aging **16**, 3137–3159 (2024).
57. Zhou, Y. et al. Human and mouse single-nucleus transcriptomics reveal TREM2-dependent and TREM2-independent cellular responses in Alzheimer's disease. Nat Med **26**, 131–142 (2020).
58. Kiani Shabestari, S. et al. Absence of microglia promotes diverse pathologies and early lethality in Alzheimer's disease mice. Cell Rep **39**, (2022).
59. Safaiyan, S. et al. White matter aging drives microglial diversity. Neuron **109**, 1100-1117.e10 (2021).
60. Park, H. et al. Single-cell RNA-sequencing identifies disease-associated oligodendrocytes in male APP NL-G-F and 5XFAD mice. Nat Commun **14**, 802 (2023).
61. Park, H. et al. Single-cell RNA-sequencing identifies disease-associated oligodendrocytes in male APP NL-G-F and 5XFAD mice. Nat Commun **14**, (2023).
62. Chen, W.-T. et al. Spatial Transcriptomics and In Situ Sequencing to Study Alzheimer's Disease. Cell **182**, 976-991.e19 (2020).
63. Mathys, H. et al. Temporal Tracking of Microglia Activation in Neurodegeneration at Single-Cell Resolution. Cell Rep **21**, 366–380 (2017).
64. Lau, S.-F. et al. IL-33-PU.1 Transcriptome Reprogramming Drives Functional State Transition and Clearance Activity of Microglia in Alzheimer's Disease. Cell Rep **31**, 107530 (2020).
65. Kristensson, K., Ghetti, B. & Wisniewski, H. M. Study on the propagation of Herpes simplex virus (type 2) into the brain after intraocular injection. Brain Res **69**, 18–201 (1974).
66. Qiu, L., Zhang, B. & Gao, Z. Lighting Up Neural Circuits by Viral Tracing. Neuroscience Bulletin vol. 38 1383–1396 Preprint at https://doi.org/10.1007/s12264-022-00860-7 (2022).
67. Oh, S. W. et al. A mesoscale connectome of the mouse brain. Nature **508**, 207–214 (2014).
68. Wickersham, I. R. et al. Monosynaptic Restriction of Transsynaptic Tracing from Single, Genetically Targeted Neurons. Neuron **53**, 639–647 (2007).
69. Ye, Q. et al. Hippocampal neural circuit connectivity alterations in an Alzheimer's disease mouse model revealed by monosynaptic rabies virus tracing. Neurobiol Dis **172**, (2022).
70. Ye, Q. et al. Monosynaptic Rabies Tracing Reveals Sex- and Age-Dependent Dorsal Subiculum Connectivity Alterations in an Alzheimer's Disease Mouse Model. Journal of Neuroscience **44**, (2024).
71. Da Silva, S. V. et al. Early synaptic deficits in the APP/PS1 mouse model of Alzheimer's disease involve neuronal adenosine A2A receptors. Nat Commun **7**, (2016).
72. Schmid, L. C. et al. Dysfunction of Somatostatin-Positive Interneurons Associated with Memory Deficits in an Alzheimer's Disease Model. Neuron **92**, 114–125 (2016).

CHAPTER 2: Exploring relationships between cellular excitability and connectivity

In this chapter, rabies-based viral tracing and computational analysis is used to explore the differences in brain connectivity that follow drug administration in mice. While the focus of this dissertation is the study of AD, and this study focused on drugs of abuse, the ventral tegmental area, and drug-induced changes in mouse brain connectivity, the insights into RABV tracing here inform experiments done in Chapter 4 and later justification for studying the RSC. We observe here that RABV tracing results are vastly different between mice treated with drugs of abuse versus saline, and that RABV counts data can differentiate between cell types in the output region.

Of particular interest in the context of our study of AD, was the identification here of a correlation between the basal expression of calcium ion channel genes to RABV count differences after drug exposure. Inhibition of expression of one such gene, *Cacna1e*, in the NAcLat reduced RABV-labeled inputs. Together, this implies that RABV count differences are indicative of a change in cellular excitability. This finding of a relationship between RABV tracing, connectivity, and cellular excitability inform our understanding of tracing experiments done in chapter 4, and, in addition to other experiments assessing excitability, justify our choice to study the retrosplenial cortex later in this work.

Chapter 2 of this dissertation was previously published in eLife. Dr. Kevin T. Beier performed the viral tracing experiments, I performed the majority of computational analyses of the data, Pieter Derdeyn assisted with network constructions and projection portraits, Dr. May Hui assisted with figure preparation, Dr. Guilian Tian, Jose J. Vasquez, Dr. Cindy M. Yamamoto, and Ghalia Azouz performed CRISPRi experiments, and Dr. Kevin T. Beier and I wrote the manuscript.

2.1 Introduction

The ventral tegmental area (VTA) contains dopamine cells (VTA^{DA}) that play key roles in several behavioral processes, including reward, aversion, and motor control[1–3]. VTA^{DA} cells are heterogeneous, with recent studies identifying subpopulations of VTA^{DA} cells that differentially contribute to adaptive and pathological behaviors[4–7]. Numerous studies have mapped the input-output architecture of VTA cells generally, as well as subpopulations of constituent DAergic, GABAergic, and glutamatergic cells[7–12]. These studies have helped us to understand the biased input and discrete output logic of the VTA, as well as generate hypotheses regarding the function of circuits including the VTA.

In addition to normal adaptive behaviors, VTA^{DA} cells critically contribute to a variety of behaviors related to substance use disorder (SUD), including reward, sensitization, and withdrawal elicited by drugs of abuse[13–18]. For a variety of drugs of abuse, even a single exposure triggers long-lasting changes in synaptic plasticity onto VTA^{DA} cells[19,20]. These changes include both a long-lasting enhancement of excitatory synaptic strength onto VTA^{DA} cells[20], and development of inhibitory synaptic plasticity[21,22]. These changes are essential for a variety of drug-induced behaviors, linking plasticity at VTA^{DA} cells with downstream behavioral changes[22–24]. Furthermore, the changes in excitatory synaptic plasticity specifically occur onto some cell types and not others[5], indicating that these adaptations are likely cell-type and pathway-specific. However, it remains unknown which sets of inputs are modified, and if these changes are consistent across abused drugs. Furthermore, synaptic plasticity is not the only type of long-lasting change that can be induced by drugs of abuse. Unfortunately, even less is known about how other forms of experience-dependent changes, including elevations in input activity, may contribute to addiction-related behaviors.

We recently showed that a single dose of cocaine causes long-lasting changes in cellular activity in inputs to VTA^{DA} cells[15,16]. These changes occur from the globus pallidus external segment (GPe) and bed nucleus of the stria terminalis (BNST) projections to the VTA that contribute to drug-induced reward and sensitization, or withdrawal anxiety, respectively[15,25]. Notably, while these studies focused on inputs from the GPe and BNST, the viral mapping results also demonstrated the existence of other input changes to the VTA that were not pursued in those studies. These unexplored connections may provide critical information about how drugs alter connectivity onto VTA^{DA} cells, which in turn contributes to drug-induced behavioral changes[15,16].

In this study, we explore the brain-wide changes in rabies-labeled inputs to VTA cells elicited by a variety of different drugs, including drugs of abuse, an anesthetic mixture of ketamine/xylazine (K/X), and fluoxetine, a psychoactive but non-abused control substance. Our study provides a first-of-its-kind exploration into how even a single drug exposure can change brain-wide connectivity patterns, in addition to highlighting some concerns about the potential effects of K/X anesthesia on motivated behavior.

2.2 Methods

Experimental Procedures

Animals: Mice were housed on a 12-hour light–dark cycle with food and water ad libitum. Males and females from a C57/BL6 background were used for all experiments in approximately equal proportions. Mice were approximately 3-4 months of age at the time of experiments. All surgeries were done under isoflurane or ketamine/xylazine anesthesia. All procedures complied with the animal care standards set forth by the National Institute of Health and were approved by the University of California, Irvine's Institutional Animal Care and Use Committee (IACUC).

Stereotaxic surgery: Mice were anesthetized with 3-4% isoflurane and maintained during surgery at 1-1.5% isoflurane, or with a mixture of ketamine (100 mg/mL) and xylazine (5 mg/mL). Mice were secured in a stereotaxic apparatus (Stoelting). Under aseptic conditions, guide holes were drilled, and viruses were infused into the target sites using a glass capillary attached to a microinjection pump (WPI, UMP3T). 100 nL of AAVs or 500 nL of EnvA-pseudotyped RABV were infused at a rate of 1.6 nL/s. The glass capillary remained in place for 5 min following the infusion to allow for virus diffusion. After infusion, the surgical incision sites were closed with either sutures or Vetbond tissue adhesive (Patterson Veterinary). Vetameg (flunixin, Patterson veterinary) was administered for pain management and topical bacitracin was applied to prevent infection at the incision site.

For viral injections mapping experiments from a defined cell type, 100 nL of a 1:1 volume mixture of AAV_5-CAG-FLEx^{loxP}-TCB and AAV_8-CAG-FLEx^{loxP}-RABV-G was injected into the VTA of DAT-Cre[26], GAD2-Cre[27], or vGluT2-Cre mice[28]. Thirteen days later, animals were given a single injection of a drug of abuse, fluoxetine, or saline. The following day, 500 nL of EnvA-pseudotyped RABV was injected into the VTA. Animals were sacrificed five days later. For cTRIO experiments shown in **Figure 2.4**, during the first surgery AAV_5-CAG-FLEx^{FRT}-TCB and AAV_8-CAG-FLEx^{FRT}-RABV-G were injected into the VTA, and CAV2-FLEx^{loxP}-Flp was injected into an output site (NAcLat, amygdala, NAcMed, mPFC) during the same surgery. All other procedures were the same as above. RABV-labeled cells throughout the brain were manually quantified, as done previously[7,10,15].

The injection coordinates and volumes used were as follows:

VTA: 100 nL AAV and 500 nL RABV, AP –3.2, LM 0.4, DV –4.2

NAcMed: 250 nL, AP +1.78, L0.4, DV –4.1

NAcLat: 250 nL, AP +1.45, LM 1.75, DV –4.0

Amygdala: 500 nL, AP –1.43, LM 2.5, DV –4.5

mPFC: 1 mL into mPFC (two injections of 500 nL, one at AP +2.15, LM 0.27, DV –2.1 and another at AP +2.15, L0.27, DV –1.6

The titers of the viruses used, based on quantitative PCR analysis, were as follows:

AAV_5-CAG-FLEx^{loxP}-TC, 2.4 x 10^{13} genome copies (gc)/mL;

AAV_8-CAG-FLEx^{loxP}-G, 1.0 x 10^{12} gc/mL;

AAV_5-CAG-FLEx^{FRT}-TC, 2.6 x 10^{12} gc/mL;

AAV_8-CAG-FLEx^{FRT}-G, 1.3 x 10^{12} gc/mL;

CAV-FLEx^{loxP}-Flp, 5.0 x 10^{12} gc/mL.

The titer of RABVDG was estimated to be 5.0 x 10^8 colony forming units (cfu)/mL based on serial dilutions of the virus stock followed by infection of the 293-TVA800 cell line.

Drug administration: Cocaine was administered (intraperitoneal injections) at a dose of 15 mg/kg, morphine at 10 mg/kg, methamphetamine at 2 mg/kg, amphetamine at 10 mg/kg, nicotine at 0.5 mg/kg, and fluoxetine at 10 mg/kg.

Quantification of RABV data: RABV mapping data were analyzed as described previously[7,10]. All RABV mapping data included were previously published[7,10,15], except for input tracing from mice treated with a single dose of methamphetamine. A total of 118 samples were included in this study (**Table 2.1**). Exclusion criteria included mice in which fewer than 200 input cells were labeled, or in cases where injections were mistargeted (e.g., starter cells in the substantia nigra pars compacta). Experimenters were blinded to the experimental condition when data quantification was being performed. For starter cell center of mass calculations, we mapped the

distribution of starter neurons, as defined by TVA-mCherry expression along the A-P axis, and in the sections at the center point along this axis, defined the center of labeling in along the M-L axis.

CRISPRi experiments

Plasmid construction: The CRISPRi system was adapted from Kumar's method[29]. pJEP317 pAAV-U6SaCas9gRNA(SapI)-EFS-GFP-KASH-pA was obtained from Addgene and was used as a template to clone pJEP317-pAAV-U6-Cacna1eAC-EFS-GFP-KASH-pA using the Q5 Site-directed mutagenesis kit (primers: 5'-acctccgtaaGTTTTAGTACTCTGGAAACAG-3' and 5'-atgaaccttGGTGTTTCGTCCTTTCCAC-3') to insert the gRNA sequence (AAGGGTTCATACCTCCGTAA) for *Cacna1e*. The gRNA sequence was selected from the IDT Predesigned Alt-R™ CRISPR-Cas9 guide RNA resource (https://www.idtdna.com/site/order/designtool/index/CRISPR_PREDESIGN). The plasmid for CRISPRi (pJEP309-pAAV-EFS-dSaCas9-KRAB-Dio-pA), pAAV2-8, and pAAV2-5 were obtained from Addgene. pHelper plasmid was a gift from Matthew Banghart at the University of California, San Diego.

AAV production, purification, and titering: Generation of AAVs from plasmids and AAV Purification by Iodixanol Gradient Ultracentrifugation were carried out as described on Addgene's website with minor modifications. Briefly, pHelper, pAAV2-5/pAAV2-8 and pJEP309/pJEP317-pAAV-U6-Cacna1eAC-EFS-GFP-KASH-pA plasmids were transfected into HEK293T cells using PEIMAX (Polysciences, cat# 24765-100). The molar ratio for pHelper:pAAV2-5:transfer plasmid was 1:1:1, and the ratio of total plasmids to PEIMAX was 1:3. 72 hrs post-transfection, cell lysate and medium precipitated with PEG were purified by iodixanol gradient. Purified AAVs were titered using primers against viral ITRs (5'-GGAACCCCTAGTGATGGAGTT-3' and 5'-CGGCCTCAGTGAGCGA-3').

Viral injection: Prior to surgery, DAT-Cre mice were anesthetized with isoflurane (4% induction, 1-1.5% maintenance) and secured on a stereotaxic frame for precise localization of injection sites. For assessment of *in vivo* genome editing, two conditions were generated. For the experimental condition, 600 nL of a 1:1:1 volume mix of AAV_8-EFS-dSaCas9-KRAB-Dio, AAV_9-CamKII-0.4-Cre-SV40 (Addgene, cat# 05558-AAV9) and AAV_5-U6-Cacna1eAC-EFS-GFP-KASH was injected into the NAcLat. For a no gRNA control condition, 400 nL of a 1:1 volume mix of AAV_8-EFS-dSaCas9-KRAB-Dio and AAV_9-CamKII-0.4-Cre-SV40 was infused into the NAcLat.

For mapping inputs to VTA cells in mice with *Cacna1e* downregulation in NAcLat, during a single surgery, 300 nL of a of a 1:1 volume mix of AAV_5-FLEx^{loxP}-TC and AAV_5-FLEx^{loxP}-RABV-G was injected into the VTA bilaterally, 600 nL of a 1:1:1 volume mix of AAV_8-EFS-dSaCas9-KRAB-Dio, AAV_9-CamKII-0.4-Cre-SV40, and AAV_5-U6-Cacna1eAC-EFS-GFP-KASH was injected into the left NAcLat, and 400 nL of a 1:1 volume mix of AAV_8-EFS-dSaCas9-KRAB-Dio and AAV_9-CamKII-0.4-Cre-SV40 was injected into the right NAcLat. This study design enabled controls within the same mice to prevent systematic biases due to differences in treatment of the mice. After 14 days, RABVDG was injected into the VTA. Mice were sacrificed five days later, and brains processed as described below.

Brain section preparation, imaging, and cell counting: Five days after RABV injection, mice underwent transcardial perfusion with 1x PBS followed by 4% formaldehyde in 1x PBS. Brains were extracted and post-fixed in 4% formaldehyde for 24 hours, and then dehydrated with 30% sucrose in 1x PBS. Brains were cut into 60 µm slices using a cryostat and mounted on Superfrost Plus™ microscope slides (Fisher, cat # 1255015). Once the brain sections were fully dried on the slide, they were rehydrated in 1xPBS containing DAPI at a dilution of 1:1000

(ThermoFisher, cat# D1306). After 10 minutes, the liquid was removed, and Fluoromount-G (SouthernBiotech, cat# 0100-01) was applied to the sections. Coverslips (Thermo Scientific, cat# 152460) were used to cover the brain sections on the slides. Images of the entire slides were captured using an Olympus IX83 microscope, and manual cell counting was performed on the acquired images.

Determination of CRISPRi-mediated knockdown efficiency: To determine the efficacy of CRISPRi knockdown of *Cacna1e* expression levels, 10 days following AAV injections into the NAcLat, NAcLat brain tissue was dissected out from the brains of injected mice, and tissue was pooled (5 mice for *Cacna1e* gRNA and 3 mice for no gRNA control). Total RNA was extracted with TRIzol, and 50 ng of total RNA was used for RT-qPCR using the Luna Universal One-Step RT-qPCR Kit (New England Biolabs, cat# E3005L) following the manufacturer's instructions. Phosphoglycerate Kinase 1 (PGK1) was used as the housekeeping gene for normalization of *Cacna1e* mRNA levels.

Dimensionality reduction of RABV input data

For RABV input mapping experiments, quantified inputs cells were binned into 22 regions: Cortex (anterior cortex), NAcMed, NAcLat, NAcCore, DStr, VP, PO, Septum, BNST, EAM, EP, GPe, PVH, LHb, MHb, CeA, LH, ZI, DR, LDT, PBN, and DCN. Altogether, this dataset consists of the input cell counts for these 22 regions across 118 mouse brains, with n=4 or 5 for each condition.

RABV input tracing data were reported as the percentage of total quantified labeled RABV inputs, as reported previously[7,10]. These data were further scaled using Z-score normalization. Principal Component Analysis (PCA) was then implemented on these normalized data using the scikit learn library in Python with default parameters[30]. Weights for PCA models were visualized

as heatmaps, reflecting the contribution of each feature to each principal component. To test for differences between groups visualized in PC space, we assessed the value for each brain for each PC and if comparing 2 conditions, used an unpaired t-test, and if >2 conditions, a one-way ANOVA and if $p < 0.05$, followed by pairwise t-tests with multiple comparisons corrections (Dunnett or Tukey, as appropriate). Input values were the first 3 principal components, and significant differences between conditions for each principal component (drugs versus controls, for example) were denoted as having a p-value of less than 0.05. Results are reported where relevant in the text or in Table S2 (available to download on eLife).

Uniform Manifold Approximation and Projection (UMAP), implemented using the UMAP library in Python, was used to generate alternative visualizations of the RABV input data[31]. To choose and standardize our UMAP parameters across analyses, we investigated which methods placed mouse brains of the same type and treatment close together most consistently. For the distance metric parameter, the metrics we found to perform best were Euclidean and cosine, but we chose to use the Euclidean metric for consistency with previous analyses[32]. For the number of neighbors parameter, we chose a large number, ⅓ the size of each of our datasets, for a more global view of the data. Other parameters were kept at the default values. To account for the stochasticity of individual embeddings, we applied UMAP 20 times to each data set and found the results to be very similar, as found in previous analyses[32]. The results were then presented as correlograms. Clustered heatmaps and plots of these data were made using a combination of the Python packages seaborn[33] and matplotlib[34]. The seaborn clustermap function was implemented with default values, including Euclidean distance as the distance metric.

Each PCA and UMAP representation of data show each animal as a point, and each point is colored by the relevant condition, either genotype or drug administered. To better visualize the

distribution of a given condition in PCA and UMAP space, we overlaid the points with ellipsoids. Ellipsoids were centered at the average coordinate of a condition and stretched one standard deviation along the primary and secondary axes.

Brain Region Selection in the Allen Mouse Brain Atlas

To compare our data with connectivity data from the Allen Institute, we converted our 22 brain regions into regions or groups of regions defined by the Allen Atlas. This was necessary as our region annotations defined while doing cell counts of the RABV data were at times different from the Allen Mouse Brain Atlas-defined regions. 22 brain regions were used for RABV analyses, but these were reduced to 20 brain regions (NAcMed, NAcLat, and NAcCore were merged into the NAc) for comparison with Allen Atlas data. In several cases, multiple Allen-defined regions were used to constitute a single region that we defined, as in the case of the cortex. A summary of these conversions is as follows:

Table 2.1. Brain region names and abbreviations used throughout chapter 2.

Full brain region name	Standard abbreviation	Allen abbreviations
Cortex	Anterior cortex	PL, ILA, ORB, GU, AI, MOs, MOp
Nucleus accumbens (medial)	NAcMed	ACB
Nucleus accumbens (lateral)	NAcLat	ACB
Nucleus accumbens (core)	NAcCore	ACB
Dorsal striatum	DStr	STRd
Ventral pallidum	VP	PALv
Preoptic area	PO	MPO, LPO
Septum	Septum	LSX, MS
Bed nucleus of the stria terminalis	BNST	BST
Extended amygdala area	EAM	MEA, AAA, NLOT, COA, MA, BMA, EPv, NDB
Internal globus pallidus	EP	GPi
External globus pallidus	GPe	GPe
Paraventricular nucleus of the hypothalamus	PVH	PVH

Lateral habenula	LHb	LH
Medial habenula	MHb	MH
Central nucleus of the amygdala	CeA	CEA
Lateral hypothalamus	LH	LHA
Zona incerta	ZI	ZI
Dorsal raphe nucleus	DR	DR, PAG
Laterodorsal tegmental nucleus	LDT	LDT
Parabrachial nucleus	PBN	PB
Deep cerebellar nuclei	DCN	CBN

It is important to note that these conversions are approximate, and brain-region specific results should be interpreted with caution, as there are known differences between atlases used in this work[35].

Allen Mouse Brain Connectivity Atlas projection portrait construction

Projection portraits of the VTA were made using data from the Allen Mouse Brain Connectivity Atlas[36], as described previously[32]. Briefly, data from 20 input brain regions, as defined in the above table, were downloaded and used for analyses. Cre-based and non-Cre-based experiments representative of each brain region were chosen for visualization. Labeled cell bodies at the site of injection often spread into adjacent brain regions or only covered a part of a region. Several experiments were chosen to account for this variation and give a better estimate for each region's projections. Experiment IDs are recorded in Table S1 in our previous work[32]. Connectivity of brain regions of interest to the VTA was visualized by showing a representative slice of the VTA and coloring the image based on high or low connectivity according to averaged AAV tracing data. AAV tracing data were averaged across experiments and across brain regions that grouped together according to weights assigned by PCA, as shown in **Figures 2.1, 2.3**, and **2.4**.

Data for projection portraits and for subsequent network construction was accessed using the python Allen SDK package, developed and maintained by the Allen Institute for Brain Science.

Gene Expression Data Access and Processing

Gene expression data was downloaded from the Allen Institute's Anatomic Gene Expression Atlas (AGEA)[37], an *in situ* hybridization dataset which has annotated expression of 19,932 genes to every brain region defined in the Allen Mouse Brain Atlas. We used code written by Eli Cornblath (https://github.com/ejcorn/mouse_abi_tool) to access and download gene expression data for our regions of interest, as done in a previous study[38]. To ensure only high-quality data were included in our analyses, we filtered experiments based on metrics used in previous analyses of this dataset[39]. We retained genes for which there were data from coronal sections, or where the Pearson correlation coefficient across experiments was over 0.5 for a particular gene in our regions of interest. This filtering left us with 4,283 unique genes for the 36 Allen-defined brain regions to be aggregated to compare with our 22 quantified input sites from the RABV mapping data. We averaged data among brain regions that were defined differently in the atlas versus our RABV quantifications, which resulted in a total of 20 brain regions for comparative analyses, as detailed above. Data were normalized using a scaled sigmoid function, a method that is robust to outliers and was used in previous analyses of these data[40].

Gene expression data analysis and comparison to RABV data

To investigate how basal gene expression might be correlated to connectivity changes after drug exposure, we used processed data from the AGEA. For each experimental condition, the RABV data for the drug-treated animals were compared to those of the saline-treated controls. To quantify this comparison, we computed a ratio, as used in **Figures 2.6 and S2.7-2.9**, by dividing

the average RABV counts in each brain region for the experimental condition by the average RABV counts in each brain region for the saline condition. This resulted in a vector of 22 ratios for each experimental condition, one ratio for each brain region. We also computed a difference, as used in **Figures 2.5 and S2.4-S2.6**, by subtracting the average RABV counts in each brain region for the experimental condition by the average RABV counts in each brain region for the saline control condition. This resulted in a vector of 22 differences for each drug comparison, one difference for each brain region.

We calculated the correlation of the condition-control differences with gene expression data for each region. The Spearman correlation coefficient between the gene expression across regions and our RABV differences was used because the correlative relationship between these values cannot be assumed to be linear. A similar method was used in previous analyses[39]. A high positive correlation between a gene and an experimental condition ratio was interpreted as a greater likelihood of a brain region to exhibit RABV labeling changes after drug exposure if this gene is highly expressed before drug exposure. A negative correlation between a gene and an experimental condition ratio was taken to mean this region would be more likely to exhibit RABV labeling changes after drug exposure if the gene expression level is low.

We performed GO analysis on the top 50 most highly correlated genes and on the top 50 most negatively correlated genes. GO analysis was performed using the EnrichR package[41]. Lists of genes with highly positive or negative correlations to our condition differences were compared to GO annotations using the coding language R. The 2021 version of two databases were queried: GO Biological Process and GO Molecular Function. GO enrichment analysis results were visualized in dot plots using the package ggplot2, as shown in **Figures 2.5** and **S2.4-S2.6**. The

'overlap' term on the x-axis of these plots means the number of genes in the correlated gene lists that were also within the list of genes for the corresponding GO term.

We next compared the drug/isoflurane ratios for a group of addictive drugs to gene expression of subsets of genes of interest (**Figure 2.6, S2.7**). Similar analyses were done for ketamine and fluoxetine (**Figure S2.8, S2.9**). The subgroups of genes used were synapse-related genes[42], ion-channel related gene[43], neurotransmitter receptor-related genes[44], differentially expressed genes related to substance dependence GWAS[45], and exo- and endocytosis-related genes[46]. Other subsets of genes of interest were defined using the Guide to Pharmacology database[47]. Correlation coefficients (Spearman's ρ) and p-values are reported in the figures.

2.3 Results

Dimensionality reduction methods identify differences in inputs of VTA cell populations

We and others previously showed that DAergic, GABAergic, and glutamatergic cells in the VTA receive inputs from similar brain regions, with some quantitative differences that varied depending on the study[10,12]. In our work we also showed that connectivity differences in the VTA are largely a consequence of spatial organization of cells within the VTA[7,10,32]. The fact that the distributions of these three VTA cell populations are largely spatially overlapping may explain why the global inputs to each population are similar. However, one limitation in interpreting these studies is that they focused on individual comparisons of input counts on a per region basis. While this is a common approach in analyzing viral tracing data, and useful for both hypothesis generation and testing, it fails to account for the higher-level organization of input patterns. Our goal here is to go beyond considering single brain regions and better understand the holistic connectivity patterns of different cell populations or drug treatment conditions.

Our first goal was to explore the potential differences in input patterns generated using RABV tracing to each of the DAergic, GABAergic, and glutamatergic neuron populations in the VTA using principal component analysis, PCA. A schematic of the experiment is shown in **Figure 2.1A**, with representative images of brain regions of interest in **Figure 2.1B**. A standard representation of these data, a bar plot, is shown in **Figure 2.1C**. PCA embedding resulted in the three cell populations clustering into largely distinct groups (**Figure 2.1E**). We considered the first two principal components (PCs), which explain the most variance in the data. In PC space, DAT-Cre brains were characterized by low PC1 and low PC2 values, GAD2-Cre brains by high PC1 values, and vGluT2-Cre brains by high PC2 values. The differences we noted were statistically significant: PC1 values of GAD2-Cre compared to DAT-Cre brains were significantly different (One-way ANOVA $p = 0.010$, DAT-Cre vs. GAD2-Cre multiple comparisons adjusted $p = 0.0087$; **Figure 2.1F**) and PC2 values of DAT-Cre compared to vGluT2-Cre brains were significantly different (One-way ANOVA $p = 0.016$, DAT-Cre vs. vGluT2-Cre multiple comparisons adjusted $p = 0.016$; **Figure 2.1G**). An advantage of PCA over other dimensionality reduction approaches is that, because it is a linear approach, the weights of each feature's contribution to each PC can be clearly interpreted. In this case, we can infer which brain regions contribute most to the positive and negative dimensions of each PC. For example, in our data, DAT-Cre and GAD2-Cre brains were separated along PC1. Brain regions that contributed to the positive dimension of PC1 were the BNST, entopeduncular nucleus (EP), GPe, central amygdala (CeA), and zona incerta (ZI) (**Figure 2.1D**). These inputs were enriched onto GABAergic cells relative to DA cells (**Figure 2.1C**). To identify where in the VTA these input populations projected, we assembled projection portraits of these inputs to the VTA as we did previously[32]. This approach uses axonal projection data available from the Allen Mouse Brain Connectivity Atlas to generate spatial heatmaps that

show “hotspots” of innervation in the midbrain[36]. We first selected experiments where GFP was expressed in the input regions of interest. Pixel values representing projection density were pulled from a representative coronal slice containing the VTA for each region. For a given set of inputs of interest, we averaged together their projections to the VTA to see if the group had any common spatial trends (**Figure 2.1H**). The projection portrait of the positive PC1 inputs indicated that they preferentially project dorsally and laterally to areas surrounding the VTA (**Figure 2.1I**). In contrast, the regions that contributed most strongly to the negative dimension of PC1 were the anterior cortex, nucleus accumbens medial shell (NAcMed), ventral pallidum (VP), preoptic area (PO), septum, lateral habenula (LHb), medial habenula (MHb), and lateral hypothalamus (LH). These inputs were more highly labeled in DAT-Cre mice. The projection portrait of these inputs highlights the whole VTA and not surrounding regions (**Figure 2.1J**). While the distribution of starter cells in the VTA in DAT-Cre and GAD2-Cre mice is spatially overlapping[7], this analysis indicates that inputs with biases onto DA cells target the VTA itself while those biased onto GABAergic cells preferentially target areas dorsal and lateral to the VTA.

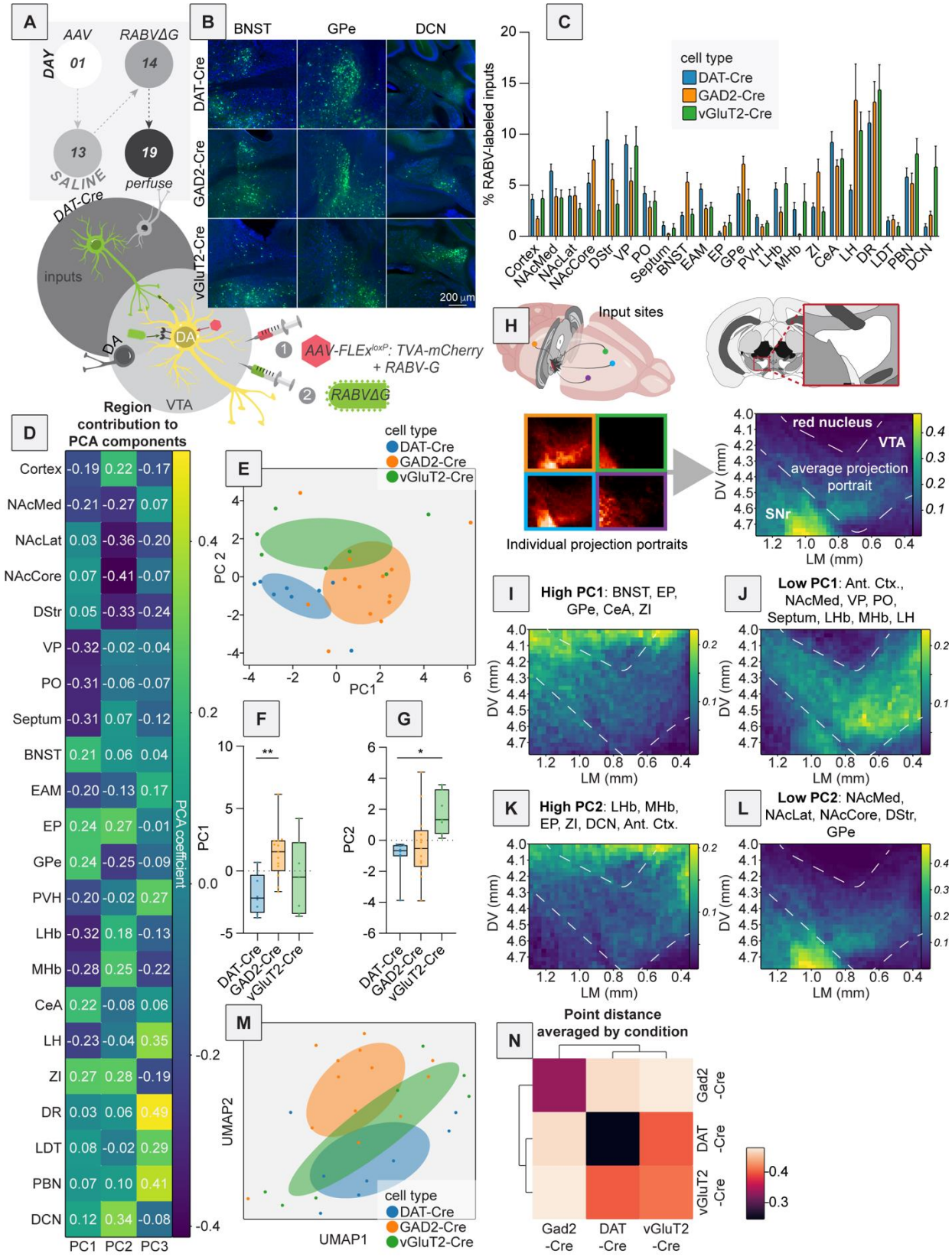
While DAT-Cre and GAD2-Cre conditions were separated along PC1, DAT-Cre and vGluT2-Cre conditions were separated along PC2. Regions that contributed to the positive dimension of PC2 included the LHb, MHb, EP, ZI, and deep cerebellar nuclei (DCN) as well as the cortex, while those contributing to the negative dimension included the NAcMed, nucleus accumbens lateral shell (NAcLat), nucleus accumbens core (NAcCore), dorsal striatum (DStr), and GPe (**Figure 2.1K-L**). The MHb and LHb both project medially in the VTA, whereas the NAcMed, NAcLat, NAcCore, DStr, and GPe all project laterally[32]. The projection portrait corresponding to the positive dimension of PC2 highlights the area dorsal of the VTA, as well as the medial aspect of the VTA (**Figure 2.1K**), while those corresponding to the negative dimension

highlight the ventrolateral aspect of the ventral midbrain, including the substantia nigra pars reticulata (SNr) (**Figure 2.1L**). These results are consistent with the medial-lateral projection gradient that principally divides the projections into the VTA, with inputs from the basal ganglia projecting to the lateral portion of the VTA and the habenula projecting to the medial portion[32]. As inputs from the MHb and LHb contribute positive weights to the PC2 axis and vGluT2-Cre brains have the most positive PC2 values, it suggests that most of the vGluT2-expressing cells in the VTA are located within the medial portion of the VTA. This is consistent with previous reports using immunohistochemistry and in situ hybridization[48,49]. In addition, the EP, ZI, and DCN all project predominantly to the red nucleus, and the cortex also innervates the red nucleus, which contains glutamatergic neurons and is located dorsal to the VTA (**Figure 2.1K**). While injections for our viral mapping studies targeted the VTA and most starter neurons were within the VTA, there were some starter neurons in vGluT2-Cre brains located in the red nucleus. In contrast, this was not the case in DAT-Cre brains[10]. As these starter populations were also largely spatially overlapping, the data indicate that inputs with biases onto glutamatergic cells target the dorsomedial aspect of the VTA, while those with a relative bias onto DA cells preferentially target the ventrolateral aspect of the VTA.

While PCA is particularly useful due to the interpretability of the weights of the PCs, it cannot capture non-linear relationships in high-dimensional data, which are common in complex systems like in neuroscience. To examine the overall relationship of input patterns between DAT-Cre, GAD2-Cre, and vGluT2-Cre cell populations, we performed a second analysis on inputs from each population using Uniform Manifold Approximation and Projection, UMAP, which can capture more complex relationships than PCA. We found that DAT-Cre and GAD2-Cre brains were the most separated in UMAP space, with vGluT2-Cre brains located between them (**Figure**

2.1M). This is consistent with the overlap in neurotransmitters released by these cells: cells expressing DA markers such as tyrosine hydroxylase (TH) and traditional GABAergic markers such as GAD2 are nearly mutually exclusive, while there is considerable overlap between cells expressing DA markers and glutamatergic markers, as well as cells expressing GABAergic markers and glutamatergic markers[10,12,50–52]. Note that while DA neurons have been found to co-transmit DA and GABA[53,54], this occurs through a non-canonical GABA synthesis pathway, and therefore DA cells do not express typical markers of GABAergic cells, such as GAD2, used to

Figure 2.1: Dimensionality reduction of RABV input tracing to 3 different VTA cell types.
(A) Schematic and timeline of viral injections into mice.
(B) Representative images of the BNST, GPe, and DCN of DAT-Cre, GAD2-Cre, and vGlut2-Cre mice. Green indicates RABV-labeled cells. Scale, 200 mm.
(C) Bar plot showing percent of RABV-labeled cells in each input region for DAT-Cre, GAD2-Cre, and vGlut2-Cre mice.
(D) Heatmap of the contributions of each brain region, or feature, in the data to PCs 1 through 3.
(E) PCA of input labeling from brains from DAT-Cre, GAD2-Cre, and vGlut2-Cre mice. For this and other figures, ellipsoids were centered at the average coordinate of a condition and stretched one standard deviation along the primary and secondary axes.
(F) Box plot comparisons of PC1. One-way ANOVA $p = 0.010$, pairwise t-tests DAT-Cre vs. GAD2-Cre multiple comparisons adjusted $p = 0.0087$, DAT-Cre vs. vGluT2-Cre $p = 0.46$, GAD2-Cre vs. vGluT2-Cre $p = 0.19$. n=8 for DAT-Cre, n = 13 for GAD2-Cre, and n = 7 for vGluT2-Cre.
(G) Box plot comparisons of PC1. One-way ANOVA $p = 0.016$, pairwise t-tests DAT-Cre vs. GAD2-Cre multiple comparisons adjusted $p = 0.65$, DAT-Cre vs. vGluT2-Cre $p = 0.016$, GAD2-Cre vs. vGluT2-Cre $p = 0.051$.
(H) Schematic of how projection portraits are constructed. A representative slice of the VTA is considered for experiments from the Allen Mouse Brain Connectivity Atlas with injections into relevant brain regions that provide input to the VTA. Projection density values in this slice are averaged and visualized to obtain the projection portrait in the VTA for that group of regions.
(I) Projection portrait of the VTA based on inputs from regions with a high positive contribution to PC1: BNST, EP, GPe, CeA, and ZI.
(J) Projection portrait of the VTA based on inputs from regions with a large negative contribution to PC1: anterior cortex, NAcMed, VP, PO, septum, LHb, MHb, and LH.
(K) Projection portrait of the VTA based on inputs from regions with a high positive contribution to PC2: LHb, MHb, EP, ZI, DCN, and anterior cortex.
(L) Projection portrait of the VTA based on inputs from regions with a large negative contribution to PC2: NAcMed, NAcLat, NAcCore, DStr, and GPe.
(M) UMAP embedding of input labeling of brains from DAT-Cre, GAD2-Cre, and vGlut2-Cre mice.
(N) Correlogram generated using the average Euclidean distance of data points from each other from 20 UMAP embeddings.


A
AAV
RABVΔG
DAY
01
14
13
19
SALINE
perfuse
DAT-Cre
inputs
DA
VTA
1 AAV-FLEx^loxP: TVA-mCherry + RABV-G
2 RABVΔG
B
BNST
GPe
DCN
DAT-Cre
GAD2-Cre
vGluT2-Cre
200 µm
C
% RABV-labeled inputs
cell type
DAT-Cre
GAD2-Cre
vGluT2-Cre
Cortex NAcMed NAcLat NAcCore DStr VP PO Septum BNST EAM EP GPe PVH LHb MHb ZI CeA LH DR LDT PBN DCN
H
Input sites
Individual projection portraits
red nucleus
VTA
average projection portrait
SNr
DV (mm)
LM (mm)
D
Region contribution to PCA components
Cortex -0.19 0.22 -0.17
NAcMed -0.21 -0.27 0.07
NAcLat 0.03 -0.36 -0.20
NAcCore 0.07 -0.41 -0.07
DStr 0.05 -0.33 -0.24
VP -0.32 -0.02 -0.04
PO -0.31 -0.06 -0.07
Septum -0.31 0.07 -0.12
BNST 0.21 0.06 0.04
EAM -0.20 -0.13 0.17
EP 0.24 0.27 -0.01
GPe 0.24 -0.25 -0.09
PVH -0.20 -0.02 0.27
LHb -0.32 0.18 -0.13
MHb -0.28 0.25 -0.22
CeA 0.22 -0.08 0.06
LH -0.23 -0.04 0.35
ZI 0.27 0.28 -0.19
DR 0.03 0.06 0.49
LDT 0.08 -0.02 0.29
PBN 0.07 0.10 0.41
DCN 0.12 0.34 -0.08
PC1 PC2 PC3
PCA coefficient
E
cell type
DAT-Cre
GAD2-Cre
vGluT2-Cre
PC 2
PC1
F
PC1
DAT-Cre GAD2-Cre vGluT2-Cre
G
PC2
DAT-Cre GAD2-Cre vGluT2-Cre
I
High PC1: BNST, EP, GPe, CeA, ZI
J
Low PC1: Ant. Ctx., NAcMed, VP, PO, Septum, LHb, MHb, LH
K
High PC2: LHb, MHb, EP, ZI, DCN, Ant. Ctx.
L
Low PC2: NAcMed, NAcLat, NAcCore, DStr, GPe
M
UMAP2
UMAP1
cell type
DAT-Cre
GAD2-Cre
vGluT2-Cre
N
Point distance averaged by condition
Gad2 -Cre
DAT -Cre
vGluT2 -Cre

define the Cre line in our study.

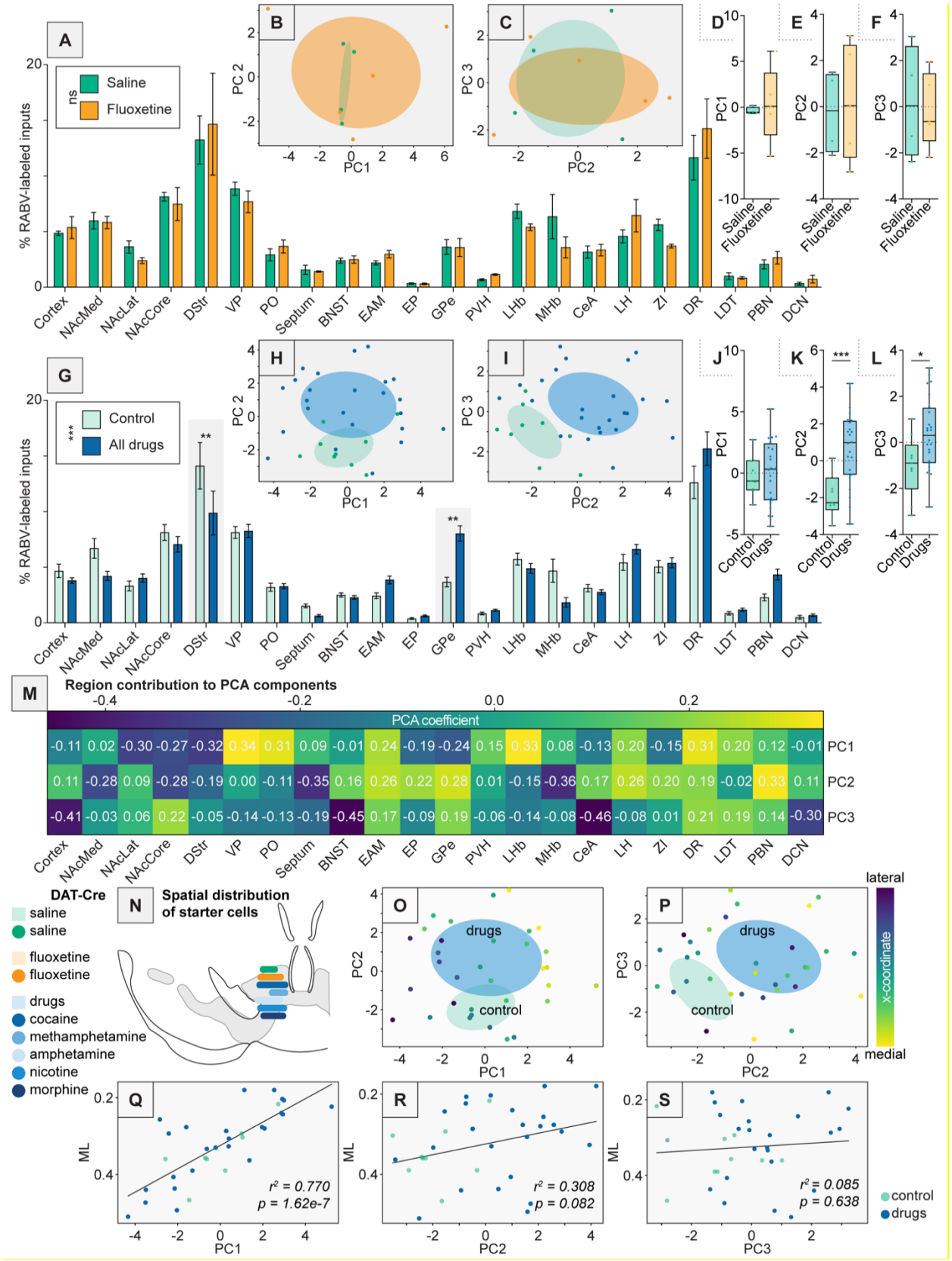
Since individual UMAP embeddings are stochastic depending on initial seeding conditions, we examined the average Euclidian distance between pairs of brains in 20 UMAP embeddings and plotted the averaged results as a correlogram. This showed that DAT-Cre and vGluT2-Cre brains were more like one another, and the GAD2-Cre brains were the most distinct from the other conditions (**Figure 2.1N**). These results are thus consistent with the known neurochemical overlap between these cells and indicate that DAT-Cre and vGluT2-Cre label more overlapping cells with each other than with GAD2-Cre. In sum, these results serve to demonstrate that we can recapitulate the locations of different neurochemically-defined cells in the VTA based on PCA and projection portraits based on information from RABV input mapping data. They also provide a validation of our approach for dissecting the relationships between input patterns and different cell types within the VTA, as the conclusions about the relationship between input regions and cell types agree with other studies using different, complementary methodologies.

A single injection of a drug of abuse changes brain-wide input patterns to VTA^{DA} cells

With our dimensionality reduction approaches validated on our cell type-specific RABV data, we next applied them to examining how drugs of abuse may alter VTA input patterns. We first assessed whether the psychoactive but non-addictive fluoxetine altered brain-wide input patterns relative to saline-injected mice. We found no clear differences in inputs between these conditions, and they overlapped completely in PC space – with no significant differences between PC1, PC2, or PC3 between the saline and the fluoxetine brains ($p > 0.05$ for PC1-3) - indicating fluoxetine did not alter brain-wide input patterns to VTA^{DA} cells (**Figure 2.2A-F**). We therefore combined these groups for additional statistical power and tested whether a single administration of one of several drugs of abuse caused changes in brain-wide input patterns to VTA^{DA} cells, first

considered as a single group. When comparing combined controls to the drugs group, we observed a statistically significant difference in input patterns, including a reduction in inputs from the DStr and elevation in inputs from the GPe following drug injection (**Figure 2.2G**). In order to more fully assess which input sites may differentiate the control- versus drug-treated groups, we again used PCA. We observed separation between the control and drug groups principally along PC2, with no separation along PC1 and minimal separation along PC3 (**Figure 2.2H-I**). There was no

Figure 2.2: PCA can deconvolve drug-induced effects from experimental variation.
(A) Bar graph representation of inputs from DAT-Cre mice anesthetized with isoflurane and treated with a single injection of saline or fluoxetine. N = 4 for saline, n = 5 for fluoxetine. Two-way ANOVA, p = 0.91.
(B) PCA plot showing PC1 and PC2 of brains from saline or fluoxetine-treated mice.
(C) PCA plot showing PC2 and PC3 of brains from saline or fluoxetine-treated mice.
(D) Box plot of PC1, p = 0.75. n = 4 for saline, n = 5 for fluoxetine.
(E) Box plot of PC2, p = 0.78.
(F) Box plot of PC3, p = 0.81.
(G) Bar graph representation of inputs from DAT-Cre mice anesthetized with isoflurane and treated with a single injection of saline or fluoxetine (controls), or one of the drugs of abuse cocaine, methamphetamine, amphetamine, nicotine, or morphine. n = 9 for control, n = 24 for combined drugs of abuse. Two-way ANOVA, p = 0.0001; unpaired t-tests with Sidak corrections for multiple comparisons: DStr, p = 0.0051, GPe, p = 0.0036.
(H) PCA plot showing PC1 and PC2 of brains from control or drug-treated mice.
(I) PCA plot showing PC2 and PC3 of brains from control or drug-treated mice.
(J) Box plot of PC1, p = 0.78. n = 9 for control, n = 24 for drugs.
(K) Box plot of PC2, p = 0.0009.
(L) Box plot of PC3, p = 0.024.
(M) Heatmap of the contributions of each brain region to PC1 through PC3 for data shown in panels E-F.
(N) Distribution of the starter cell center of mass along the medial-lateral axis for each experimental condition. Saline n = 4, fluoxetine n = 5, cocaine n = 5, methamphetamine n = 4, amphetamine n = 5, nicotine n = 5, morphine n = 5.
(O) PC1 vs. PC2 for control vs. drug conditions as shown in E, with each point colored according to the medial-lateral coordinate of the center of mass of starter cells for each brain.
(P) PC2 vs. PC3 for control vs. drug conditions as shown in F, with each point colored according to the medial-lateral coordinate of the center of mass of starter cells for each brain.
(Q) Linear regression between the x-coordinate of the center of mass of starter cells for each brain vs. PC1. $r^2 = 0.770$, p = 1.62e-7,
(R) Linear regression between the x-coordinate of the center of mass of starter cells for each brain vs. PC2. $r^2 = 0.308$, p = 0.082.
(S) Linear regression between the x-coordinate of the center of mass of starter cells for each brain vs. PC3, $r^2 = 0.085$, p = 0.638.


A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
Saline
Fluoxetine
ns
% RABV-labeled inputs
Control
All drugs
***
**
*
PC1
PC2
PC3
Cortex
NAcMed
NAcLat
NAcCore
DStr
VP
PO
Septum
BNST
EAM
EP
GPe
PVH
LHb
MHb
CeA
LH
ZI
DR
LDT
PBN
DCN
Region contribution to PCA components
PCA coefficient
DAT-Cre
saline
fluoxetine
drugs
cocaine
methamphetamine
amphetamine
nicotine
morphine
Spatial distribution of starter cells
lateral
medial
x-coordinate
control
ML
r² = 0.770
p = 1.62e-7
r² = 0.308
p = 0.082
r² = 0.085
p = 0.638

significant difference between drugs and controls in PC1 ($p = 0.78$; **Figure 2.2J**). The differences in PC2 and PC3 were statistically significant: as we observed visually, the control group's PC2 and PC3 values were significantly different from the drug group's PC2 and PC3 values ($p = 0.0009$ and $p = 0.024$ for PC2 and PC3, respectively; **Figure 2.2K-L**). The positive axis of PC2, which was enriched in drug-treated brains, was driven by inputs from EAM, EP, GPe, LH, and PBN, while the negative axis was driven by the NAcMed, NAcCore, DStr, septum, and MHb (**Figure 2.2M**). Notably, the most negative weights for PC2, namely the NAcMed, NAcCore, DStr, septum, and MHb are all reduced in drug-treated brains (**Figure 2.2G**). Conversely, the most positive weights, namely the EAM, EP, GPe, LH, and PBN, are all increased in drug-treated brains (**Figure 2.2G**). Therefore, while PCA allows us to observe trends consistent with traditional methods, we can also observe the relationship between the percentage of inputs among different brains, and also identify relationships that are not immediately apparent in the data when presented as bar graphs. For example, while the BNST, CeA, and ZI do not show clear differences in control vs. drug-treated brains in the bar graphs, they are all positive weights in PC2 and thus, likely change in concert with regions such as the GPe and PBN (**Figure 2.2G-M**).

Importantly, unlike traditional bar graph approaches, PCA allows us to isolate technical variation from the experiments, enabling us to focus more clearly on changes induced by the relevant variable, here treatment with a control substance or drug of abuse. For example, PC1 does not distinguish control vs. drug-treated brains (**Figure 2.2H, J**), indicating this variable was inherent to the way the experiment was performed and not related to drug or control treatment. The strongest negative weights for PC1 were the NAcLat, NAcCore, DStr, and GPe, while the strongest positive weights for PC1 were the VP, PO, LHb, and DR (**Figure 2.2M**). This is consistent with connectivity differences along the medial-lateral axis of the VTA, which underlies

many of the connectivity differences to different VTA^{DA} subpopulations[7,10,32]. To assess whether PC1 may reflect differences the location of starter cells in the VTA, we first assessed whether there was a systematic difference in starter cell location between conditions. We noted a high level of overlap in starter cell distribution between conditions (**Figure 2.2N**), indicating that this alone likely did not explain differences between control and drug-treated groups. We then assessed whether brain to brain variation in starter cell distribution could explain the effects observed along PC1. We therefore colored each point in PC space according to the assessed center of injection for each experiment. We noted a clear medial-lateral gradient along PC1, but not PC2 or PC3 (**Figure 2.2O-P**). To test the significance of this relationship, we performed linear regression analysis, comparing the x-coordinate of the injection center relative to location along PC1, PC2, or PC3 for each brain. We observed a highly significant positive correlation with PC1, a slightly positive but non-significant correlation with PC2, and no correlation for PC3 (**Figure 2.2Q-S**). These results show that PC1 reflects variation in the injection site within the VTA, while PC2 reflects changes induced by drugs of abuse, allowing us to isolate the changes in input connectivity induced by administration of a single drug of abuse.

With our integrative methodology validated on cell type specific RABV data and demonstrating that drugs of abuse alter brain-wide input patterns to VTA^{DA} cells, we applied our approaches to explore the effects of each drug specifically, as well as potential effects of the 2x use of the K/X anesthetic mixture. We first asked two questions: 1) How do injections of different drugs of abuse change the connectivity patterns onto VTA cells? 2) How do these potential drug-induced differences relate to differences between unique cell types in the VTA? To answer these, we first assessed experiments with ten separate experimental conditions: 2 conditions of GAD2-Cre mice, and 8 conditions of DAT-Cre mice. The GAD2-Cre mice were all anesthetized with

isoflurane and administered either saline or cocaine one day prior to RABV injection. The DAT-Cre mice were given either isoflurane or K/X-based anesthesia. The K/X-anesthetized mice were administered saline. The isoflurane anesthetized mice were administered either saline, fluoxetine, amphetamine, cocaine, methamphetamine, morphine, or nicotine. RABV mapping was performed from targeted VTA cells in all cases, and the brain-wide inputs were quantified. We then performed PCA on the combined RABV tracing datasets, which showed that brains from GAD2-Cre mice were in the positive dimension of PC1, while those from DAT-Cre mice were in the negative dimension (**Figure 2.3A**), regardless of the substance administered. This difference in PC1 was highly statistically significant (One-way ANOVA $p < 0.0001$; all conditions pairwise comparisons with GAD2-Cre, multiple comparisons adjusted $p < 0.0001$; **Figure 2.3B**). The positive dimension of PC1 received contributions from largely the same sets of regions as in the previous analysis of DAT-Cre, GAD2-Cre, and vGluT2-Cre brains, which included the BNST, EP, ZI, CeA, as well as the DCN (**Figure 2.3H, 2.1D**). Given that the same set of brain regions differentiated DAT-Cre and GAD2-Cre brains along PC1 in both datasets, this suggests that the intrinsic differences in inputs to different cell populations in the VTA are more pronounced than differences induced by drugs of abuse, fluoxetine, or the method of anesthesia. Notably, these patterns were unlikely to be obtained by chance; scrambling the sample identity and percentage of RABV-labeled cells in each input site resulted in groups that overlapped in both PCA and UMAP space (**Figure S2.1**).

To specifically explore potential drug-induced differences in input labeling onto VTA^{DA} cells, we removed GAD2-Cre brains from the dataset. Results and representative images are shown in **Figure 2.3J-K**. In PCA analyses of these brains only, PC1 and PC3 largely failed to differentiate the conditions, with no significant differences observed (**Figure 2.3C-G**), while PC2 significantly differentiated the conditions (One-way ANOVA $p = 0.0093$; saline vs. drugs multiple comparisons

adjusted p = 0.034; **Figure 2.3F**). The most negative weights to PC1 were contributed by the basal ganglia inputs that project to the SNr, located ventrolaterally to the VTA (**Figure 2.3I, L**) while positive contributors came from brain regions that projected more medially (LHb, MHb, dorsal raphe (DR)) and centrally to the VTA (ventral pallidum (VP), PO, extended amygdala area (EAM), paraventricular nucleus of the hypothalamus (PVH), and LH) (**Figure 2.3I, M**). This is consistent with a gradient along the medial-lateral axis[32]. As it did not differentiate the conditions, this axis likely represents experimental variations in the location of starter cell populations within the VTA[7,10,32]. The regions with the largest positive contribution to PC2 included the EAM, parabrachial nucleus (PBN), DR, EP, and LH, with lesser contributions from the GPe, DCN, ZI, CeA, and BNST (**Figure 2.3I, N**). These regions were over-represented in animals treated with a single administration of a drug of abuse. Notably, this group of inputs contains brain regions that are involved in the brain's stress response such as the PBN, EAM, CeA, and BNST, and have been implicated in drug-induced behaviors[55,56]. For example, we recently showed the importance of the GPe in conditioned place preference and locomotor sensitization and the BNST in drug-induced withdrawal[15,25]. The regions with the largest negative contribution to PC2 included the NAcMed, NAcCore, DStr, Septum, and MHb (**Figure 2.3I, O**). These regions were underrepresented in animals given a single administration of a drug of abuse. These results together demonstrate that drugs of abuse cause long-lasting changes in input populations to DA cells, and that we can identify which sets of inputs distinguish drug-treated vs. control-treated brains. These changes may be functionally important, as we previously demonstrated the functional role of two inputs that we identified as differing between drug-treated and control conditions.

Lastly, we wanted to explore the overall relationship between input changes induced by different drugs. As before, we used UMAP to explore the higher-level organization of these

Figure 2.3: Dimensionality reduction analysis of brains from animals treated with a single drug exposure.

(A) PCA plot showing PC1 and PC2 of brains from DAT-Cre and GAD2-Cre mice. DAT-Cre mice were anesthetized with K/X and treated with saline or anesthetized with isoflurane and treated with saline or one of the drugs cocaine, methamphetamine, amphetamine, nicotine, morphine, or fluoxetine. GAD2-Cre mice were anesthetized with isoflurane and treated with saline or cocaine.

(B) Box plot of PC1, One-way ANOVA $p < 0.0001$, pairwise t-tests DAT-Cre saline vs. DAT-Cre fluoxetine multiple comparisons adjusted $p = 0.99$, DAT-Cre saline vs. DAT-Cre K/X $p = 1.0$, DAT-Cre saline vs. DAT-Cre drugs $p = 0.23$, DAT-Cre saline vs. GAD2-Cre $p < 0.0001$, DAT-Cre fluoxetine vs. DAT-Cre K/X $p = 1.0$, DAT-Cre fluoxetine vs. DAT-Cre drugs $p = 0.46$, DAT-Cre fluoxetine vs. GAD2-Cre $p < 0.0001$, DAT-Cre K/X vs. DAT-Cre drugs $p = 0.31$, DAT-Cre K/X vs. GAD2-Cre $p < 0.0001$, DAT-Cre drugs vs. GAD2-Cre $p < 0.0001$. $n = 4$ for DAT-Cre saline and DAT-Cre K/X, $n = 5$ for DAT-Cre fluoxetine, $n = 24$ for DAT-Cre drugs, $n = 17$ for GAD2-Cre.

(C) PCA plot showing PC1 and PC2 of brains from DAT-Cre mice anesthetized with isoflurane and treated with cocaine, methamphetamine, amphetamine, nicotine, morphine, fluoxetine, or saline, or anesthetized with K/X and treated with saline.

(D) PC2 and PC3 of the same group of brains as shown in panel B.

(E) Box plot of PC1, One-way ANOVA $p = 0.81$.

(F) Box plot of PC2, One-way ANOVA $p = 0.0093$, pairwise t-tests DAT-Cre saline vs. DAT-Cre fluoxetine multiple comparisons adjusted $p = 0.94$, DAT-Cre saline vs. DAT-Cre K/X $p = 0.87$, DAT-Cre saline vs. DAT-Cre drugs $p = 0.034$, DAT-Cre fluoxetine vs. DAT-Cre K/X $p = 0.99$, DAT-Cre fluoxetine vs. DAT-Cre drugs $p = 0.10$, DAT-Cre K/X vs. DAT-Cre drugs $p = 0.27$.

(G) Box plot of PC3, One-way ANOVA $p = 0.086$.

(H) Heatmap of the contributions of each brain region to PC1 through PC3 for data shown in panel A.

(I) Heatmap of the contributions of each brain region to PC1 through PC3 for data shown in panels B and C.

(J) Bar plot showing percent of RABV-labeled cells in each input region for the DAT-Cre mice shown in panel B.

(K) Representative images of brain slices from DAT-Cre mice showing, from top to bottom, the NAc, BNST, GPe, and PBN. Scale, 500 mm.

(L) Projection portrait of the VTA based on inputs from regions with a large negative contribution to PC1, based on data from panel E: NAcCore, NAcLat, DStr, GPe.

(M) Projection portrait based on inputs from regions with a high positive contribution to PC1, based on data from panel E: LHb, DR, VP, PO, EAM, LH.

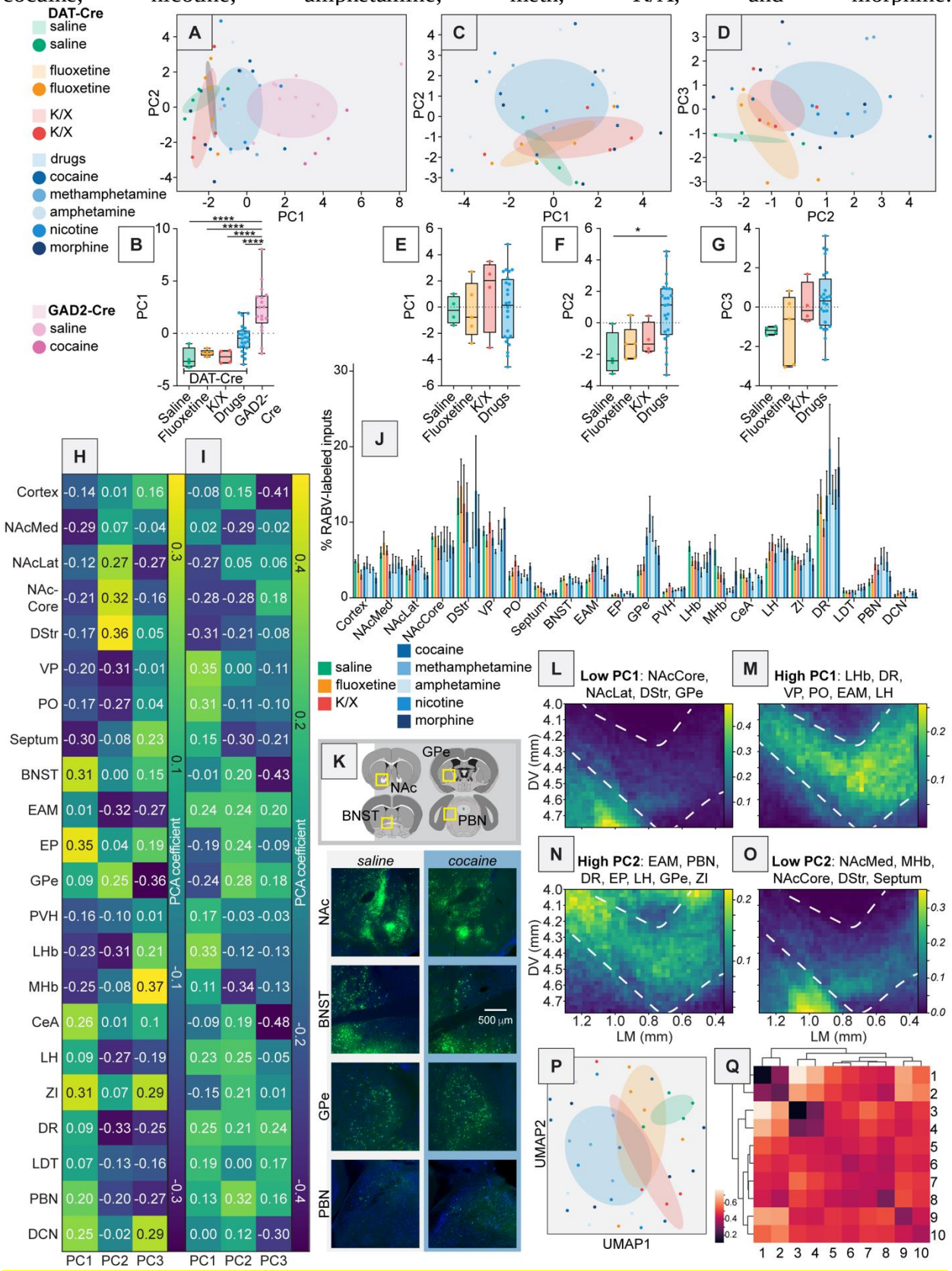
(N) Projection portrait based on inputs from regions with a high positive contribution to PC2, based on data from panel E: EAM, PBN, DR, EP, LH, GPe, ZI.

(O) Projection portrait based on inputs from regions with a large negative contribution to PC2, based on data from panel E: NAcMed, NAcCore, DStr, Septum, MHb.

(P) UMAP embedding of data from the brains of DAT-Cre mice shown in panels B and F.

(Q) Correlogram showing average Euclidean distance of data points over 20 UMAP embeddings of the data presented in panel J. Experiment conditions from bottom to top and left to right are labelled 1-10 and are ordered as follows: GAD2-Cre cocaine, GAD2-Cre saline, saline, fluoxetine,

cocaine, nicotine, amphetamine, meth, K/X, and morphine.



conditions to compare the relationship of holistic input patterns between groups. We found that the mice anesthetized with isoflurane and treated with saline overlapped with the isoflurane-anesthetized, fluoxetine-treated group, but not the K/X-anesthetized and saline-treated group (**Figure 2.3P**). These conditions were segregated from animals treated with the drugs of abuse - psychostimulants, morphine, and nicotine. Our analysis of the relationships between individual conditions, generated using the Euclidian distances between points in 20 UMAP embeddings as before, demonstrated that the expected control conditions formed a cluster (saline, fluoxetine) distinct from the other groups. Interestingly, the K/X-anesthetized animals were more closely associated with the mice treated with addictive drugs, in particular morphine, than with controls (**Figure 2.3Q**). This suggests that drugs of abuse cause similar changes in brain-wide input patterns to one another in DAT-Cre mice, and that input patterns in K/X-anesthetized animals more closely resemble morphine-treated mice than saline-treated mice. This indicates that the method of anesthesia, either using isoflurane or K/X, results in different patterns of input labeling. Given the similarity of input patterns in K/X-anesthetized mice to isoflurane-anesthetized mice treated with a drug of abuse but not saline, we suspect that two doses of K/X-anesthesia as is required for RABV circuit mapping experiments triggers changes in input patterns in rodent brains that last at least the five-day duration of these experiments.

Anesthetic doses of K/X cause long-lasting changes in inputs to $VTA^{DA}\rightarrow$NAcLat and $VTA^{DA}\rightarrow$Amygdala cells

To examine how differences in anesthesia methods may impact brain-wide input patterns, we performed some more targeted comparisons. When comparing only two groups using PCA – isoflurane-anesthetized animals to K/X-anesthetized animals – these conditions completely segregated (**Figure 2.4A**). While the difference in PC1 was not quite significant ($p = 0.086$), the

difference in PC2 was significant (p = 0.037). This separation was driven by positive weights from the VP, PO, EAM, PVH, LH, and PBN along PC1, which were more strongly represented in the K/X-anesthetized mice and whose combined projection portrait broadly targets the VTA (**Figure 2.4A, C, P, U**), and the NAcLat, DStr, GPe, and PVH along PC2, which were also more strongly represented in the K/X-anesthetized mice whose combined projection portrait targets ventrolateral to the VTA (**Figure 2.4A, C, P, V**). A similar separation of conditions was observed using UMAP (**Figure 2.4B**).

Next, we wanted to assess which VTA^{DA} cells may be impacted by whether the animals received isoflurane or K/X-based anesthesia. To test this, we performed cell type-specific Tracing of the Relationship between Inputs and Outputs (cTRIO) for each of four different VTA^{DA} subpopulations: those projecting to the NAcMed, NAcLat, amygdala, or mPFC, and receiving K/X anesthesia and saline, isoflurane and saline, or isoflurane and cocaine. The differences in input patterns between conditions were most pronounced for $VTA^{DA}\rightarrow$NAcLat cells, where the isoflurane- and K/X-treated groups segregated, both in PCA and UMAP space (**Figure 2.4D-F**). Groups appeared to differ along both PC1 and PC2, though differences in PC2 were not statistically significantly different, with K/X-treated brains being characterized by lower PC1 values which received contributions principally from the anterior cortex, NAcMed, NAcLat, NAcCore, septum, PVH, and MHb, as well as higher PC2 values which received contributions from the VP, PO, EAM, LHb, and LH (PC1, One-way ANOVA p = 0.0003, saline vs. cocaine multiple comparisons adjusted p = 0.011, k/x vs. cocaine multiple comparisons adjusted p = 0.0002; **Figure 2.4D, Q, W, X**). The following comparisons in PC1 were statistically significant: cocaine versus isoflurane controls (p = 0.011) and cocaine versus K/X anesthesia (p = 0.001). Notably, K/X, saline-treated mice and isoflurane, cocaine-treated mice segregate from one another in both PCA and UMAP

Figure 2.4: Dimensionality reduction analysis of brains from animals anesthetized with isoflurane and treated with either cocaine or saline, or anesthetized with ketamine/xylazine and treated with saline.

(A) PCA plot of input data from VTA^{DA} cells from mice anesthetized either with isoflurane or K/X and given saline. Box plot comparisons of PC1 and PC2 are shown to the right for this plot and D, G, J, and M. PC1, p = 0.086, PC2, p = 0.037, unpaired t-tests. N = 4 for both.

(B) UMAP plot of input data from VTA^{DA} cells from mice anesthetized either with isoflurane or K/X and given saline.

(C) Bar plot showing percent of RABV-labeled cells in each input region from mice anesthetized either with isoflurane or K/X and given saline.

(D) PCA plot of RABV input mapping experiments to $VTA^{DA}\rightarrow$NAcLat cells in mice anesthetized with isoflurane and given saline or cocaine, or anesthetized with K/X and given saline. PC1, One-way ANOVA p = 0.0003, pairwise t-tests saline vs. cocaine multiple comparisons adjusted p = 0.011, saline vs. K/X p = 0.071, K/X vs. cocaine p = 0.0002. PC2, One-way ANOVA p = 0.16. n=4 for saline and K/X, n = 5 for cocaine.

(E) UMAP plot of input mapping experiments to $VTA^{DA}\rightarrow$NAcLat cells.

(F) Bar plot showing percent of RABV-labeled cells in each input region for cTRIO experiments from $VTA^{DA}\rightarrow$NAcLat cells.

(G) PCA plot of input mapping experiments to $VTA^{DA}\rightarrow$Amygdala cells. PC1, One-way ANOVA p = 0.48. PC2, One-way ANOVA p = 0.30. n=4 for saline, n = 5 for K/X and cocaine.

(H) UMAP plot of input mapping experiments to $VTA^{DA}\rightarrow$Amygdala cells.

(I) Bar plot showing percent of RABV-labeled cells in each input region for cTRIO experiments from $VTA^{DA}\rightarrow$Amygdala cells.

(J) PCA plot of input mapping experiments to $VTA^{DA}\rightarrow$NAcMed cells. PC1, One-way ANOVA p = 0.64. PC2, One-way ANOVA p = 0.27. n=4 for saline and K/X, n = 5 for cocaine.

(K) UMAP plot of input mapping experiments to $VTA^{DA}\rightarrow$NAcMed cells.

(L) Bar plot showing percent of RABV-labeled cells in each input region for cTRIO experiments from $VTA^{DA}\rightarrow$NAcMed cells.

(M) PCA plot of input mapping experiments to $VTA^{DA}\rightarrow$mPFC cells. PC1, One-way ANOVA p = 0.67. PC2, One-way ANOVA p = 0.97. n=4 for saline and cocaine, n = 5 for K/X.

(N) UMAP plot of input mapping experiments to $VTA^{DA}\rightarrow$mPFC cells.

(O) Bar plot showing percent of RABV-labeled cells in each input region for cTRIO experiments from $VTA^{DA}\rightarrow$mPFC cells.

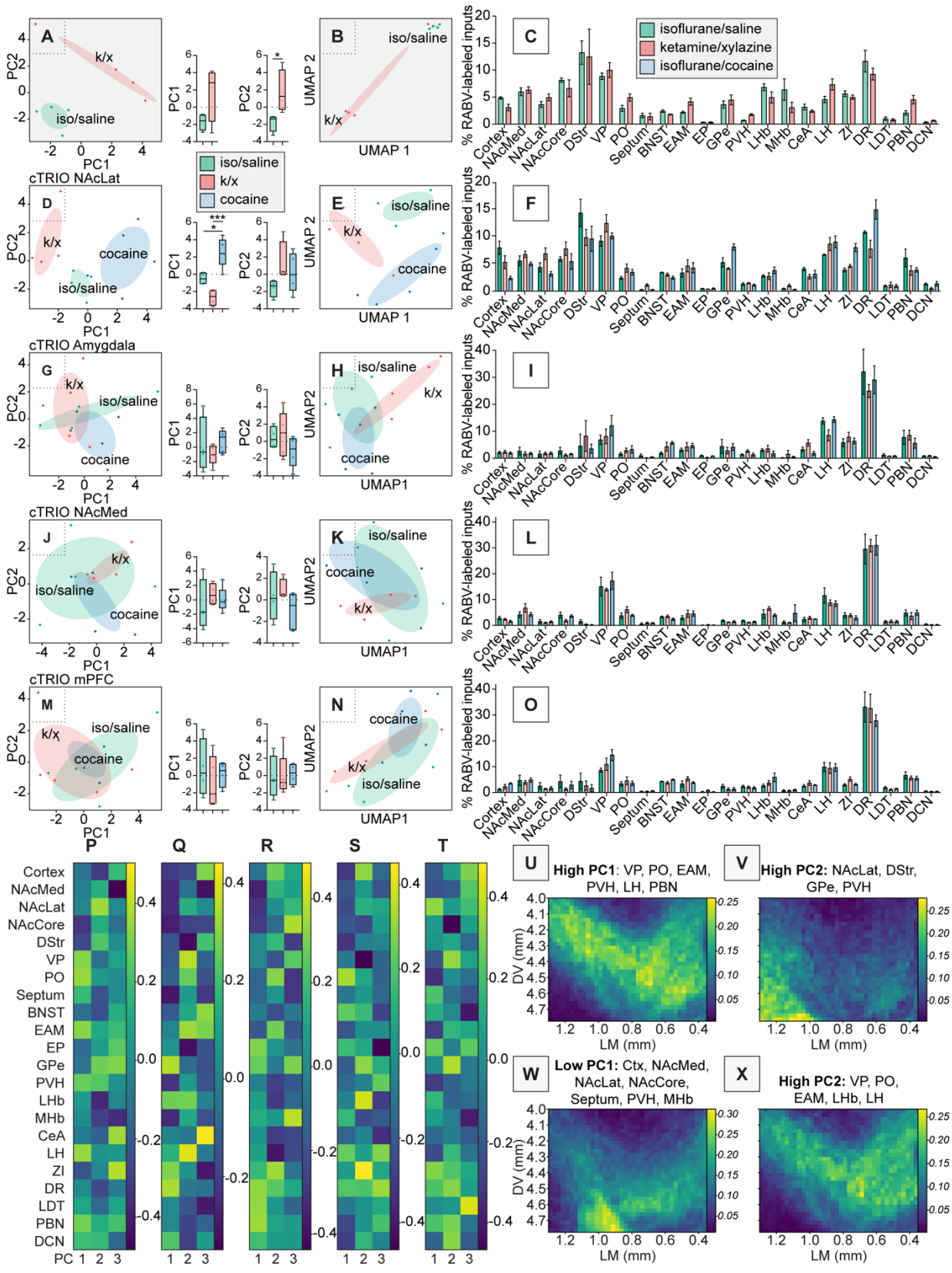
(P-T) Heatmap of the contributions of each brain region to PC1 to PC3 for data shown in (P) panel A, (Q) panel D, (R) panel G, (S) panel J, (T) panel M.

(U) Projection portrait of regions with high PC1 contributions in all VTA^{DA} cells: VP, PO, EAM, PVH, LH, PBN.

(V) Projection portrait of regions with high PC2 contributions in VTA^{DA} cells: NAcLat, DStr, GPe, PVH.

(W) Projection portrait of regions with low PC1 contributions in $VTA^{DA}\rightarrow$NAcLat cells: cortex, NAcMed, NAcLat, NAcCore, septum, PVH, MHb.

(X) Projection portrait of regions with high PC2 contributions in $VTA^{DA}\rightarrow$NAcLat cells: VP, PO, EAM, LHb, LH.


A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
cTRIO NAcLat
cTRIO Amygdala
cTRIO NAcMed
cTRIO mPFC
k/x
iso/saline
cocaine
iso/saline
ketamine/xylazine
isoflurane/cocaine
PC1
PC2
UMAP 1
UMAP1
UMAP 2
UMAP2
% RABV-labeled inputs
Cortex
NAcMed
NAcLat
NAcCore
DStr
VP
PO
Septum
BNST
EAM
EP
GPe
PVH
LHb
MHb
CeA
LH
ZI
DR
LDT
PBN
DCN
PC 1 2 3
High PC1: VP, PO, EAM, PVH, LH, PBN
High PC2: NAcLat, DStr, GPe, PVH
Low PC1: Ctx, NAcMed, NAcLat, NAcCore, Septum, PVH, MHb
High PC2: VP, PO, EAM, LHb, LH
DV (mm)
LM (mm)

space, indicating that the K/X-induced changes are distinct from those induced by cocaine (**Figure 2.4D-F**). The similarity in projection portraits for weights that contribute more to the K/X-treated group in VTA^{DA}→NAcLat (**Figure 2.4U, V**) and all VTA^{DA} neurons (**Figure 2.4W, X**) suggests that largely overlapping DA cells may be targeted for each set of experiments and is consistent with the observation that VTA^{DA}→NAcLat neurons comprise the majority of VTA^{DA} neurons, as noted in previous studies[7,10]. These results thus serve as another proof of principle of the reliability of our approach to pull out similar results from different datasets.

In addition to VTA^{DA}→NAcLat cells, we observed some segregation of isoflurane and K/X-treated conditions in VTA^{DA}→Amygdala cells (**Figure 2.4G-I, R**). The lack of clear differentiation in PCA space, with only PC3 showing significant differences between the isoflurane control group and K/X anesthesia (One-way ANOVA $p = 0.014$, saline vs. K/X multiple comparisons adjusted $p = 0.0078$; **Figure S2.2**) makes it difficult to interpret which input sites may be contributing to the differentiation in conditions that we observe. In the VTA^{DA}→NacMed and VTA^{DA}→mPFC subpopulations, the separation between isoflurane-saline, isoflurane-cocaine, and K/X-saline conditions was much less pronounced (**Figure 2.4J-O, S-T**), and not to be significant; indeed, similar separations were observed by chance (**Figure S2.3**). Notably, when considering inputs to all VTA^{DA} cells, the projection portraits that indicate the inputs most highly elevated or depressed in K/X-treated mice highlight either the entire VTA, or the ventrolateral aspect of the VTA (**Figure 2.4U, V**). These projections are consistent with the locations of the VTA^{DA}→Amygdala and VTA^{DA}→NAcLat cells, respectively, within the VTA. In contrast, the portraits do not highlight the medial or ventromedial aspects of the VTA, which is the location of VTA^{DA}→mPFC and VTA^{DA}→NAcMed cells[32,57]. These results together

indicate that K/X anesthesia impacts inputs that preferentially target $VTA^{DA}\rightarrow$NAcLat or $VTA^{DA}\rightarrow$Amygdala cells.

Exploring gene expression patterns that predict changes in RABV input labeling

We next wanted to explore if drug-induced changes in input connectivity might be related to gene expression profiles in cells located in the input sites. The AGEA data are from all cells in the input regions and are not limited to those projecting to the VTA and show only baseline gene expression patterns and not those following drug administration. Nonetheless, these analyses can be a valuable starting point in associating spatial gene expression data with patterns of RABV-labeled inputs from VTA^{DA} cells. To do this, we calculated the correlation of gene expression data from the Allen Gene Expression Atlas (AGEA) with the difference (**Figure 2.5**) or the ratio (**Figure 2.6**) between inputs labeled in each region in experimental conditions and isoflurane/saline-treated controls. After filtering AGEA data for only high-quality experiments, we were left with 4,283 genes to which we could calculate correlations within our 22 input brain regions of interest.

We first identified genes or pathways that may be linked to elevations or depressions in RABV labeling for each condition. We selected genes that were highly positively or negatively correlated to differences between 1) isoflurane-saline controls and 2) isoflurane- psychostimulant, morphine, nicotine, or fluoxetine, or 3) K/X-saline mice. We used the top 50 genes in the AGEA that were either positively or negatively correlated (Spearman's ρ) to differences in RABV labelling and performed gene ontology (GO) analyses. Starting with drugs of abuse in aggregate, the 50 most positively correlated genes (Spearman's $\rho > 0.47$, $p < 0.05$) were most significantly associated with inhibitory ligand-gated chloride channels (**Figure 2.5A**), suggesting that higher basal expression of these genes may be related to larger changes in connectivity induced by drugs

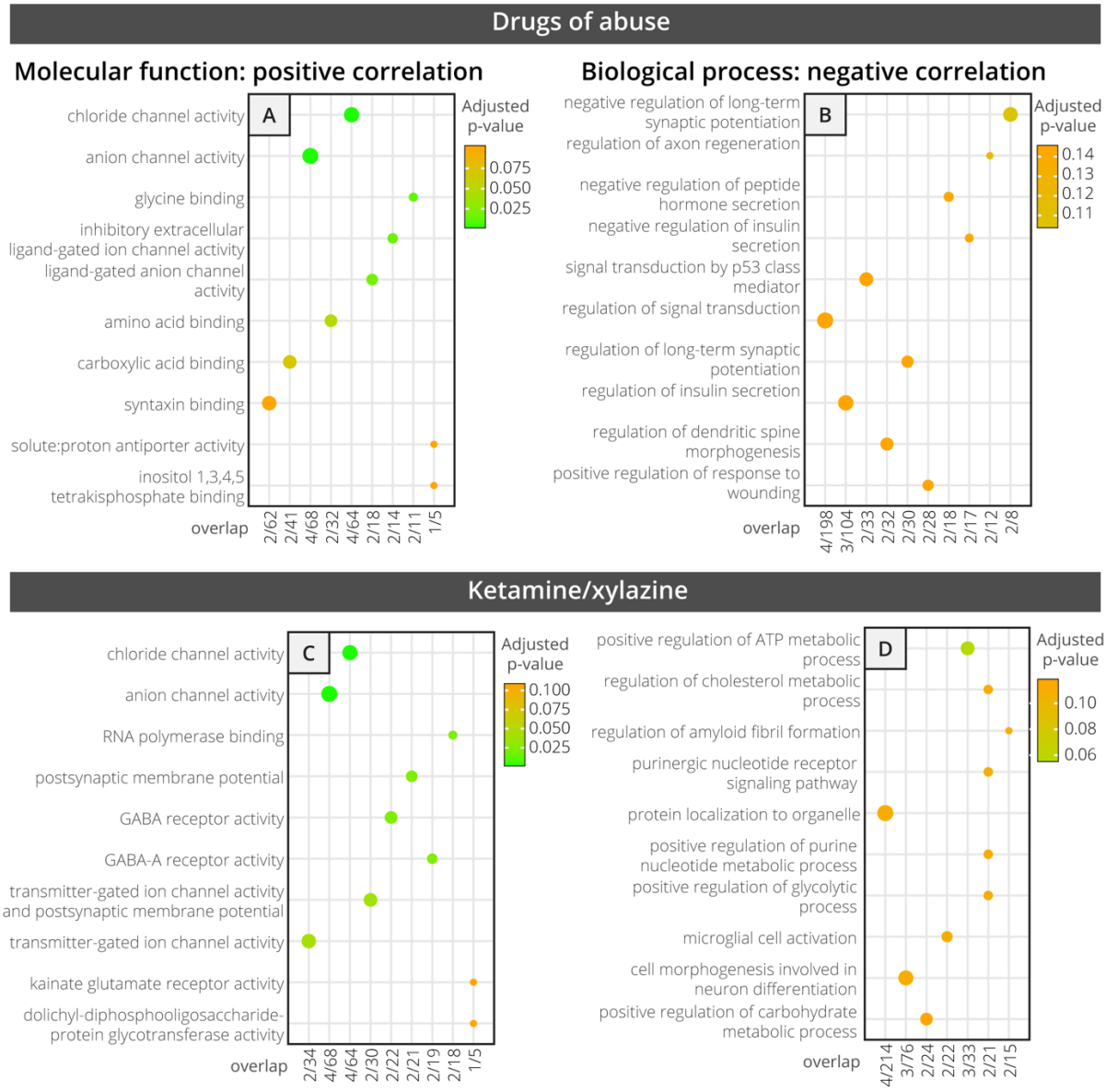


Figure 2.5: Gene ontology (GO) analysis of top 50 genes within the Allen Gene Expression Atlas correlated with differences in RABV input labeling obtained from experimental vs. control mice.

(A-B) GO from an average of the addictive drug conditions vs. saline-treated controls, all anesthetized using isoflurane.

(C-D) GO from K/X-anesthetized vs. isoflurane-anesthetized saline-treated mice.

Panels A and C show molecular function-related gene classes whose expression was positively correlated with RABV labeling differences, and panels B and D show GO analyses that reflect biological process-related gene classes whose expression was negatively correlated with RABV labeling differences.

of abuse. Conversely, there were many more genes that were strongly negatively correlated to drug-induced changes in RABV labeling. The 50 most negatively correlated genes (Spearman's $\rho < -0.68$, $p < 0.001$) were most closely associated with negative regulation of long-term synaptic potentiation (**Figure 2.5B**), suggesting that regulation of synaptic plasticity in input cell populations may be a key indicator of the likelihood of a drug of abuse altering RABV input labeling. We repeated this analysis of the top 50 genes with each individual drug: cocaine, methamphetamine, amphetamine, morphine, and nicotine. In addition to findings in aggregate, other interesting classes of genes that were negatively associated (Spearman's $\rho < -0.50$, $p < 0.05$) with individual drug-induced changes in RABV labeling include regulation of synapse structural plasticity (methamphetamine), regulation of calcium ion transport (nicotine), cellular response to Ca^{2+} ions (cocaine), and regulation of potassium ion transporters (nicotine) (**Figure S2.4**), which relate to synaptic and cellular processes of plasticity and intrinsic excitability. There were fewer than 50 genes that were significantly positively correlated, but the top 50 genes (Spearman's $\rho > 0.36$, $p < 0.12$) were associated with terms including anion channels (cocaine, amphetamine, methamphetamine) and syntaxin binding (cocaine, amphetamine, nicotine) (**Figure S2.5**). We also explored which gene expression patterns may be related to input changes driven by K/X anesthesia. Anion channel activity, particularly GABA receptor expression, again shows up as one of the most significant associations to the top 50 most positively correlated genes (Spearman's $\rho > 0.37$, $p < 0.10$) (**Figure 2.5C**), while negatively correlated genes were associated with a variety of metabolic processes (**Figure 2.5D**). Notably, these gene expression patterns were not correlated with input changes induced by fluoxetine (**Figure S2.6**).

To further assess which classes of genes were the most closely related to the differences in drug evoked RABV labeling relative to controls, we performed linear regressions of the ratios of

RABV labeling in experimental vs. control conditions relative to the mean of the log normalized expression of AGEA RNA *in situ* hybridization data for a particular class of genes. We observed an overall significant and slightly negative correlation between the RABV labeling ratio and the gene expression mean ($r = -0.30$, $p = 2.1 \times 10^{-3}$) for all 4,283 genes (**Figure S2.7A**). This genome-wide negative correlation is a useful reference to compare to correlation coefficients of subsets of genes. Among the gene expression classes that we examined, the most significant association was between the change in RABV-labeled inputs and ion channel gene expression ($r = -0.56$, $p = 1.3 \times 10^{-9}$; **Figure 2.6A**), consistent with our results from GO analyses. The next most significant classes were neurotransmitter receptors ($r = -0.46$, $p = 1.9 \times 10^{-6}$) and synapses ($r = -0.42$, $p = 1.6 \times 10^{-5}$), followed by genes that have been linked to SUDs through Genome-Wide Association Studies (GWAS; $r = -0.33$, $p = 7.7 \times 10^{-4}$) as well as endo/exocytosis genes ($r = -0.30$, $p = 2.3 \times 10^{-3}$). Each of the above relationships showed a negative correlation of RABV ratios with gene expression (**Figure 2.6B-C, S2.7B-C**). Except for endo/exocytosis genes, each class had a more negative, significant correlation than our genome-wide reference. As a further comparison, we observed that the relationships between RABV labeling ratios and gene classes that are likely irrelevant to RABV labeling, such as mitochondrial genes ($r = -0.24$, $p = 1.4 \times 10^{-2}$), sugar transporters ($r = -0.26$, $p = 1.0 \times 10^{-2}$), and tumor-associated genes ($r = -0.18$, $p = 7.0 \times 10^{-2}$) were weaker and less significant than our classes of interest (**Figure S2.7D-F**).

As the class of ion channel genes was the most strongly correlated to drug-induced differences in RABV-labeled input patterns, we assessed which ion channels are most closely linked to RABV labeling differences between conditions. The most significant relationship was between Ca^{2+} channel genes ($r = -0.57$, $p = 4.4 \times 10^{-10}$), followed by Cl^- ($r = -0.39$, $p = 6.1 \times 10^{-5}$), K^+ ($r = -0.30$, $p = 2.9 \times 10^{-3}$), and Na^+ channels ($r = -0.11$, $p = 0.30$; **Figure 2.6D-G**). The

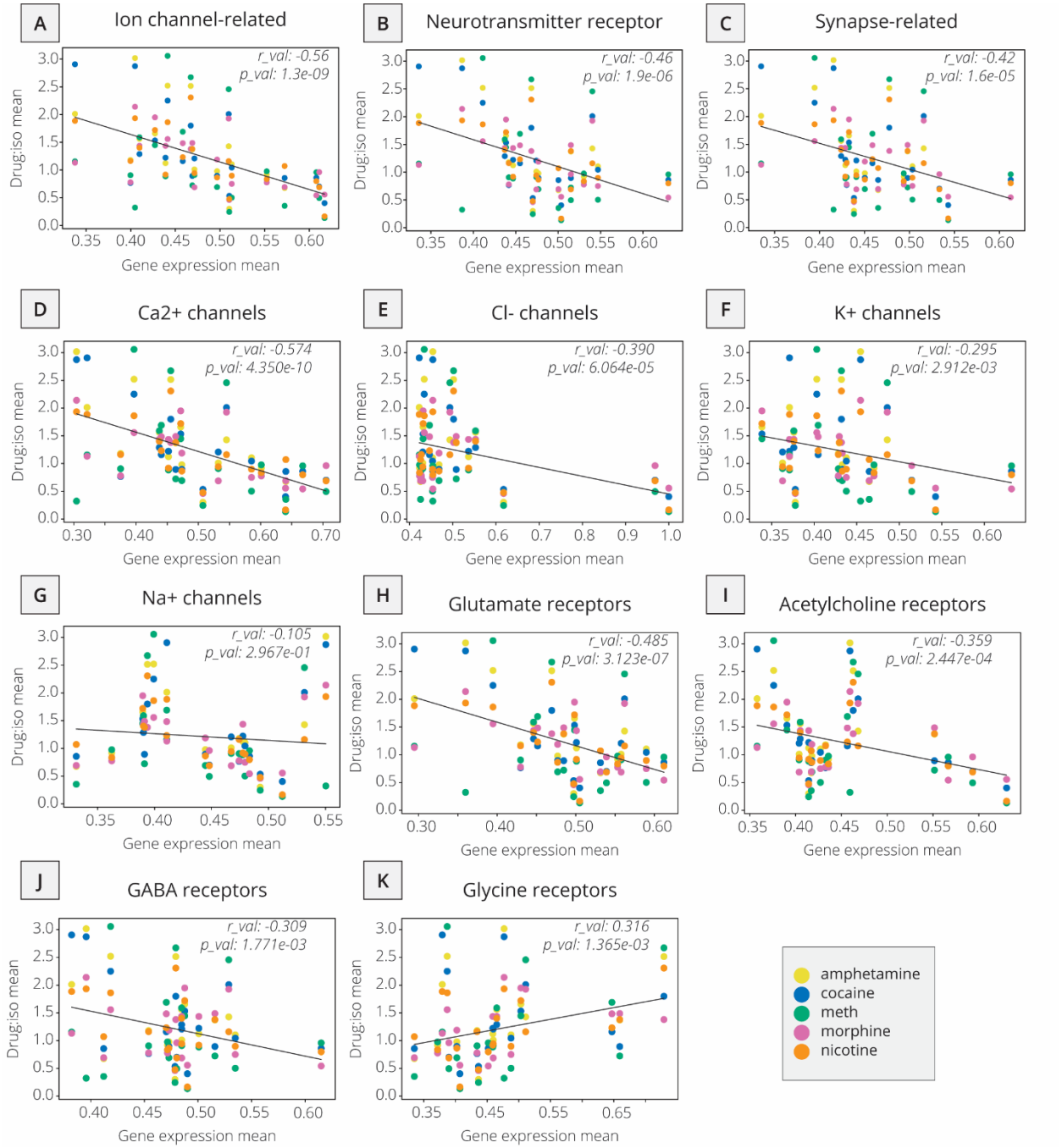


Figure 2.6: Relationship between RABV labeling ratio in addictive drug/saline-treated groups and mean basal gene expression of gene subgroups, subtypes of ion channels, and neurotransmitter receptors. Linear regressions are shown for all five examined drugs against different gene classes expressed by brain region within the Allen Gene Expression Atlas, including (A) ion channel-related genes, (B) neurotransmitter-related genes, (C) synapse-related genes, (D) Ca^{2+} channels, (E) Cl^- channels, (F) K^+ channels, (G) Na^+ channels, (H) glutamate receptors, (I) acetylcholine receptors, (J) GABA receptors, (K) glycine receptors.

expression levels of these channels, in particular Ca^{2+} channels, are linked to regulating a variety of cellular processes, including immediate early gene expression networks as well as cellular excitability. Ligand-gated channels ($r = -0.40$, $p = 4.1 \times 10^{-5}$) were more strongly related with RABV labeling than voltage-gated ion channels ($r = -0.33$, $p = 7.4 \times 10^{-4}$) or other types of ion channels ($r = -0.13$, $p = 0.21$; **Figure S2.7G-I**). Among the ligand-gated channels, the glutamate receptors were the most strongly correlated ($r = -0.49$, $p = 3.1 \times 10^{-7}$; **Figure 2.6H**). The next most strongly correlated gene classes were acetylcholine receptors ($r = -0.36$, $p = 2.4 \times 10^{-4}$), GABA receptors ($r = -0.31$, $p = 1.8 \times 10^{-3}$), and glycine receptors ($r = 0.32$, $p = 1.4 \times 10^{-3}$) (**Figure 2.6I-K**). Notably, glycine receptors were the only class of genes that we analyzed that was positively correlated with RABV labeling. Similar patterns were observed for K/X anesthetized mice (**Figure S2.8**), but not with isoflurane-anesthetized, fluoxetine-treated mice (**Figure S2.9**). These results together indicate that the basal expression levels of a variety of ion channels, particularly Ca^{2+} and ligand-gated ion channels, may be related to driving elevations or depressions in RABV labeling following a single injection of a variety of drugs of abuse, or K/X anesthesia.

As the correlation between RABV labeling ratios and Ca^{2+} channel gene expression is negative, lower expression of Ca^{2+} genes is linked to an elevation in drug-induced RABV input labeling. This relationship could be explained by several mechanisms. One, expression of Ca^{2+} channel genes could be inhibitory towards the drug-induced change in RABV transmission. Alternatively, lower Ca^{2+} channel gene expression levels could indicate more potential for a drug-induced rise in gene expression level. To differentiate between these possibilities, we performed CRISPR-based gene expression modulation of the voltage-gated calcium ion channel *Cacna1e* using CRISPRi, which reduces gene expression through blocking DNA transcription without cutting DNA[58] (**Figure 2.7A-B**). *Cacna1e* was chosen as it was one of the Ca^{2+} channel genes used

in **Figure 2.6D** that had a strong negative correlation with RABV labeling. Three different adeno-associated viruses (AAV) were co-injected into the NAcLat, an input site to the VTA, in DAT-Cre mice. These viruses expressed 1) Cre, 2) Cre-dependent dCas9, 3) Cre-dependent small guide RNA (gRNA) targeting *Cacna1e*. A two AAV mixture lacking the AAV expressing the gRNA was used as a control. The three AAV experimental mix was administered into the NAcLat in one hemisphere, while the two AAV control mix was injected into the other hemisphere. During the same surgery, AAVs expressing Cre-dependent TVA-mCherry and rabies virus glycoprotein were bilaterally injected into the VTA, so that RABV mapping could be done in both hemispheres (**Figure 2.7C-D**). Because the NAcLat, NAcCore, and GPe unilaterally project to the VTA in the same hemisphere, and not the contralateral VTA[7,10,15,25], the ratio of labeled inputs in these sites on a single hemisphere can be used to assess the relative labeling in the NAcLat in experimental and control conditions within each mouse.

We first validated that our CRISPRi strategy reduced *Cacna1e* gene expression. We designed our gRNA to cause only a mild reduction in *Cacna1e* expression to more naturally mimic potential changes in gene expression. We found that our CRISPRi perturbation reduced gene expression by ~20% (**Figure 2.7E**). This knockdown of *Cacna1e* in the NAcLat via CRISPRi resulted in decreased RABV-labeled inputs from the NAcLat compared to control mice where the shRNA was not included (**Figure 2.7F**). These results indicate that reduced levels of *Cacna1e* lower the number of RABV-labeled inputs from the NAcLat, suggesting that expression levels of *Cacna1e* are directly related to RABV-labeled inputs. This means that expression of this gene is not inhibitory towards RABV transmission. Therefore, a more likely hypothesis is that a low basal gene expression of Ca^{2+} channels allows for greater changes in the expression of these genes following drug exposure.

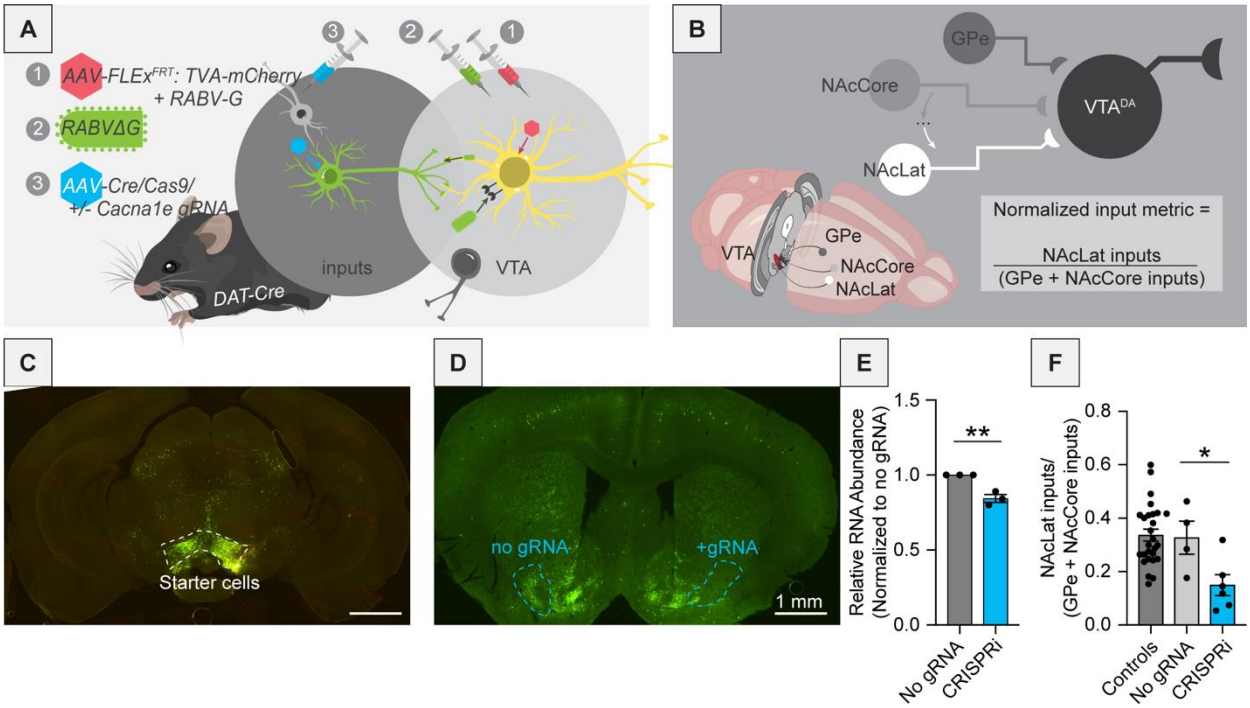


Figure 2.7: Inhibition of *Cacna1e* expression in the NAcLat reduces RABV-labeled inputs from the NAcLat onto VTA^{DA} cells.
(A) Schematic of viral injections performed.
(B) Schematic of connectivity in relevant brain regions and explanation of the normalized input metric.
(C) Representative image of fluorescence in starter cells in the VTA. Scale, 1 mm.
(D) Representative image of fluorescence in the NAcLat after either administration of no gRNA (left) or gRNA targeting *Cacna1e* (right). Scale, 1 mm.
(E) qPCR results showing slight inhibition of *Cacna1e* in the NAcLat.
(F) Difference in RABV-labeled inputs after CRISPRi-mediated knockdown of *Cacna1e* compared to controls, and no gRNA.

2.4 Discussion

Here we explored the long-term consequences of a single injection of a variety of abused drugs on the connectivity of VTA^{DA} cells. We first confirmed that dimensionality reduction approaches could adequately differentiate the connectivity between different cell populations in the VTA, and showed which inputs differentiate DAergic, GABAergic, and glutamatergic cells in the VTA (**Figure 2.1**). We then showed that a variety of drugs of abuse as well as K/X-based anesthesia trigger long-lasting changes in the brain-wide input patterns to VTA cells, which were

clearly observable though less substantial than the differences onto different cell types or natural variation in starter cell location (**Figures 2.2-2.3**). We used our cTRIO method to identify DA cell type-specific changes induced by cocaine as well as K/X anesthesia, implicating the $VTA^{DA}\rightarrow$NAcLat and $VTA^{DA}\rightarrow$Amygdala cells as those exhibiting the most substantial input changes (**Figure 2.4**). We then explored intrinsic gene expression patterns throughout the brain that relate to the elevation or depression in RABV labeling by different drugs, finding that expression levels of ion channels, neurotransmitter receptors, and synapse-related genes were all significantly negatively correlated with RABV input labeling ratios, with Ca^{2+} channels and glutamate receptors being the most strongly correlated (**Figure 2.6**). We then demonstrated the causative role of one calcium channel gene, *Cacna1e*, in influencing RABV input mapping efficiency (**Figure 2.7**).

Drugs of abuse induce changes in connectivity onto VTA^{DA} neurons from common and distributed input populations throughout the brain

Using dimensionality reduction approaches we differentiated the input patterns to different VTA cells (**Figure 2.1**) and showed that these differences were larger than the differences in inputs from the same cell populations in animals given a drug of abuse vs. saline (**Figure 2.3A**). Notably, these differences were unique from those triggered by the psychoactive but non-addictive substances fluoxetine and K/X (**Figure 2.3C, F**). Differences were apparent on the 2nd PC axis, which were driven by positive weights from the GPe, EP, EAM, LH, ZI, DR, and PBN, regions that showed elevated connectivity in drug-treated mice. Many of these regions have been implicated in cocaine addiction[59–62]. While these drug-induced differences in input labeling were less substantial than intrinsic cell type connectivity differences (**Figure 2.3A**) or differences in starter cell location (**Figure 2.2N-S**), these differences are more apparent when using

dimensionality reduction methods such as PCA and UMAP. That different cell types exhibit larger differences in input patterns than the same cell type with and without drug administration is perhaps not surprising, as different populations of cells may receive different sets of inputs from one another, and these are likely more stable across experiments than the differences in connectivity triggered following a single exposure to a drug of abuse. The same is likely true for spatial gradients, which we determined to be a major source of variation in RABV inputs to the VTA[32]. Notably, the minor differences in starter cell location are likely not a technical concern specific to our study, but rather inherent to all RABV mapping experiments that rely on intracranial injections, as a small amount of biological or technical variation is to be expected. The PCA approach we employ here, however, can better isolate these effects than traditional bar graph approaches and thus enable clearer observation of the effect of the relevant variable being tested, here effects of a single drug administration (or K/X anesthesia) on the input connectivity to VTA^{DA} cells.

The results in this study further confirm our previous observations that a single exposure to a variety of drugs of abuse triggers long-lasting changes in inputs to VTA^{DA} neurons. Our interpretation of these data is that the changes are largely being driven by a change in the activity of input cells, as we have shown using fiber photometry and slice electrophysiology in three separate studies[15,16,63]. However, we have not confirmed this to be true for all cases. As a single exposure to a variety of drugs of abuse alters the strength of both excitatory and inhibitory inputs in the VTA[5,20–22], we cannot rule out that changes in synapse number or strength may also influence the observed connectivity results.

Ketamine/xylazine anesthesia alters the input landscape of VTA^{DA} cells

Animals anesthetized with a K/X mixture showed a different pattern of input labeling to VTA^{DA} cells than animals anesthetized with isoflurane (**Figure 2.4A-B**). While we cannot perform RABV mapping on unanesthetized mice, making it impossible to definitively define which mode of anesthesia, or both, alters RABV input patterns we suspect that isoflurane anesthesia has minimal effect on the brain-wide inputs to VTA^{DA} cells. This is largely because animals treated with a drug of abuse with isoflurane anesthesia look more like those receiving K/X-saline than isoflurane-saline animals (**Figure 2.3**). In addition, ketamine is a drug with known abuse potential[64–67]. While the anesthetic doses used here (100 mg/kg) are well above the doses used recreationally (~1 mg/kg), the changes that we observed with K/X anesthesia may be related to the abuse potential of ketamine. Interestingly, the K/X-induced input changes were most similar to those induced by 10 mg/kg of morphine (**Figure 2.3P-Q**). Antidepressant doses of ketamine may work in part via the opioid system; indeed, it has been known for some time that ketamine binds the mu opioid receptor[68–71], and it was more recently shown that the opioid antagonist naltrexone blocks the antidepressant effects of ketamine[72,73]. Given that anesthetic doses of ketamine (100 mg/kg) are roughly 10 times higher than antidepressant doses of ketamine (~10 mg/kg), K/X anesthesia almost certainly engages the opioid system.

Relationship of gene expression with the extent of drug-induced RABV input labeling

Gene expression patterns in the mouse brain are correlated with the extent of elevations and depressions of RABV labeling induced by one of several drugs of abuse. We observed a slightly negative correlation of RABV labeling changes with the overall level of gene expression across the brain. We are not sure of the implications of this negative correlation, though it could suggest that if initial gene expression levels are lower, it leaves a higher potential for experience-induced elevations. We observed significant negative correlations of RABV labeling changes

induced by drugs of abuse with the levels of expression of ion channel genes, neurotransmitter receptor-related genes, and synapse-related genes. These results and our experiments demonstrating the functional impact of downregulating *Cacna1e* for RABV input labeling (**Figure 2.7**) indicate that the expression levels of these genes could be used as predictive biomarkers of which regions are likely to undergo drug-induced changes in connectivity to VTA^{DA} cells. They also suggest that the expression levels of these gene classes are related to the drug-induced changes in RABV labeling.

While the AGEA is a useful resource for relating spatial gene expression information to our RABV mapping datasets, there are notable limitations. First, the AGEA dataset is not as quantitatively rigorous as modern single nucleus RNA sequencing or spatial transcriptomic datasets. Nonetheless, we believe that the overall patterns of high and low expression, especially when considered among classes of genes as was performed here, can still be informative. Furthermore, the AGEA dataset provides a static picture of gene expression levels throughout the brain, and thus does not account for potential changes in gene expression levels triggered by administration of one of the drugs used in this study. Third, the AGEA dataset considers expression levels in all cell types in a brain region, irrespective of their individual transcriptomic/molecular identities, while only a subset of cells in a region project to VTA^{DA} cells. Nonetheless, our CRISPRi experiment showing that a mild reduction of *Cacna1e* reduces RABV labeling indicates the potential for causative associations between gene expression patterns and drug-induced RABV input labeling changes (**Figure 2.7**). This result also directly links cellular excitability and the extent to which RABV labels these cells as monosynaptic inputs to defined cell populations.

Our RABV screen identified several input sites that show elevated or depressed labeling following a single administration of a drug of abuse, suggesting potential changes in excitability

of these cells. Ca^{2+} channel genes were also highly negatively correlated with the extent of drug-induced changes in RABV input labeling. Ca^{2+} channels have been implicated in substance abuse through Genome-Wide Association Studies[74], and molecules that target these channels have shown some promise as addiction therapeutics[75–80]. It is of particular interest that Ca^{2+} and not Cl^-, Na^+, or K^+ channels are highly correlated to RABV input labeling changes. Ca^{2+} is critically involved as a secondary messenger in neurons that affects a variety of processes including long-lasting changes in transcription and synaptic transmission, whereas the other ions are mostly involved only in facilitating or inhibiting formation and propagation of the action potential. Understanding how Ca^{2+} channel gene expression levels relate to drug-induced changes in the brain will help to illuminate the relationship between basal gene expression levels, drug-induced changes in RABV labeling, and substance abuse, providing mechanistic insights into how inter-individual differences influence susceptibility to substance abuse.

2.5 Supplementary figures

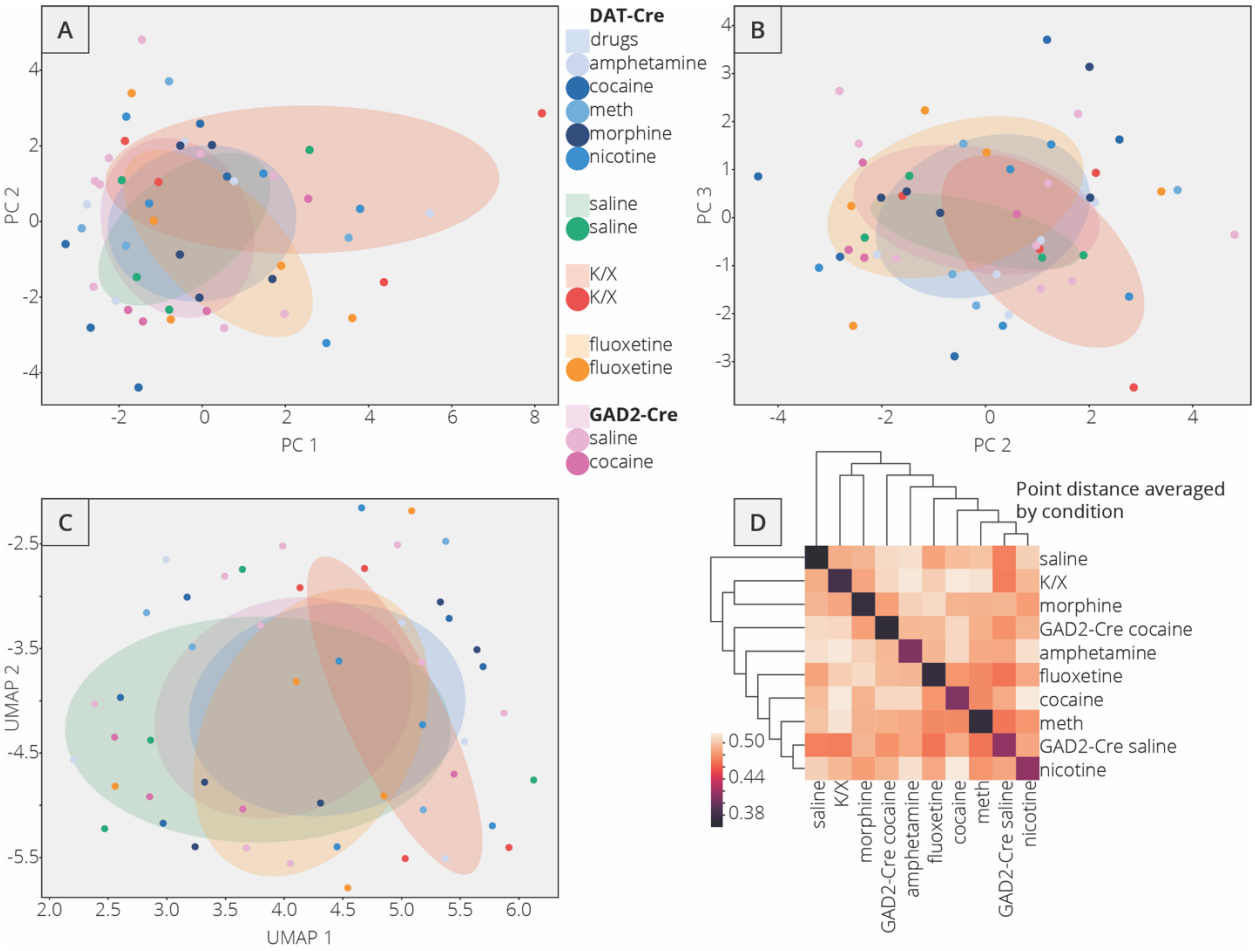


Figure S2.1: Brains from animals treated with a single exposure to a drug (Figure 2.3A), but with the brain identities scrambled.
(A) PCA plot showing the data with scrambled identities along PC1 and PC2.
(B) The same scrambled data along PC1 and PC3.
(C) UMAP of the scrambled data.
(D) Correlogram of 20 UMAP embeddings of the scrambled data.

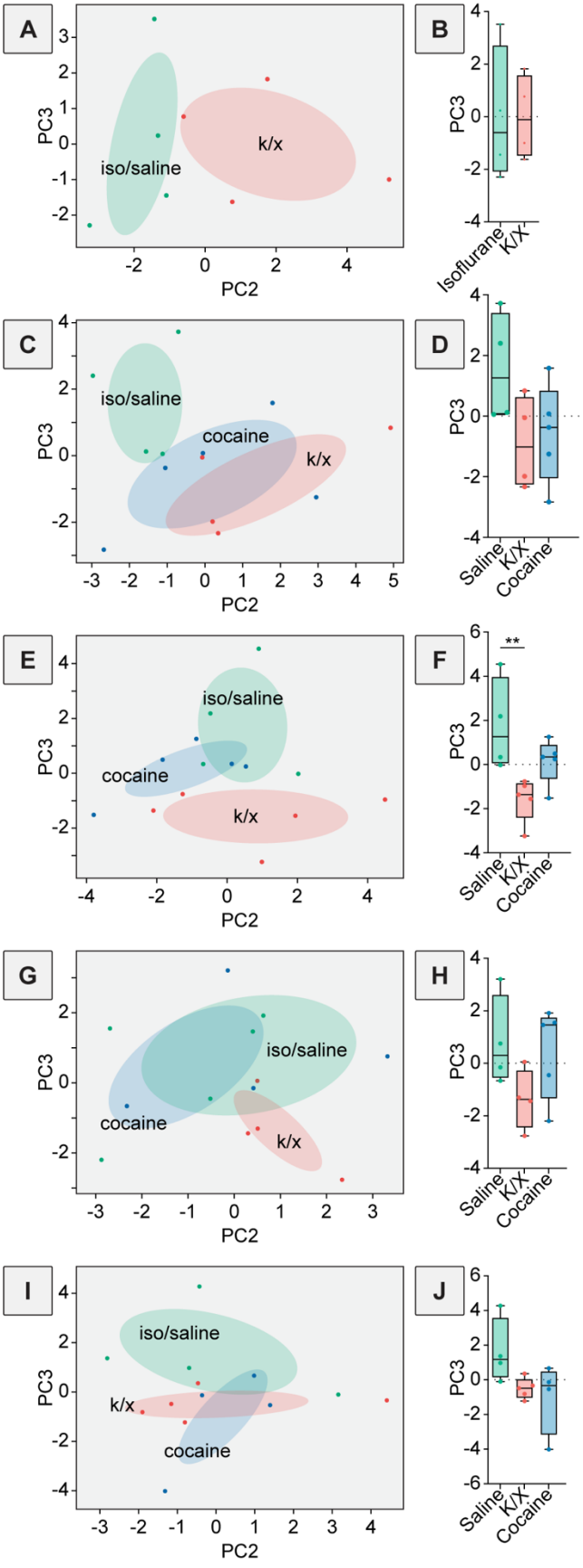


Figure S2.2: Dimensionality reduction analysis of brains from animals anesthetized with isoflurane and treated with either cocaine or saline, or anesthetized with ketamine/xylazine and treated with saline, focusing on PC3.

(A) PCA plot of input data from VTA^{DA} cells from mice anesthetized either with isoflurane or K/X and given saline, showing PC2 and PC3.

(B) Box plot comparisons of PC3, p = 0.99, unpaired t-tests. N = 4 for both.

(C) PCA plot of RABV input mapping experiments to VTA^{DA}→NAcLat cells in mice anesthetized with isoflurane and given saline or cocaine, or anesthetized with K/X and given saline, showing PC2 and PC3.

(D) Box plot comparisons of PC3, One-way ANOVA p = 0.12. n=4 for saline and K/X, n = 5 for cocaine.

(E) PCA plot of input mapping experiments to VTA^{DA}→Amygdala cells, showing PC2 and PC3.

(F) Box plot comparisons of PC3, PC1, One-way ANOVA p = 0.014, pairwise t-tests saline vs. K/X multiple comparisons adjusted p = 0.011, saline vs. cocaine p = 0.24, K/X vs. cocaine p = 0.16. n=4 for saline, n = 5 for K/X and cocaine.

(G) PCA plot of input mapping experiments to VTA^{DA}→NAcMed cells, showing PC2 and PC3.

(H) Box plot comparisons of PC3,, One-way ANOVA p = 0.64. PC2, One-way ANOVA p = 0.16. n=4 for saline and K/X, n = 5 for cocaine.

(I) PCA plot of input mapping experiments to VTA^{DA}→mPFC cells, showing PC2 and PC3.

(J) Box plot comparisons of PC3, One-way ANOVA p = 0.67. PC2, One-way ANOVA p = 0.083. n=4 for saline and cocaine, n = 5 for K/X.

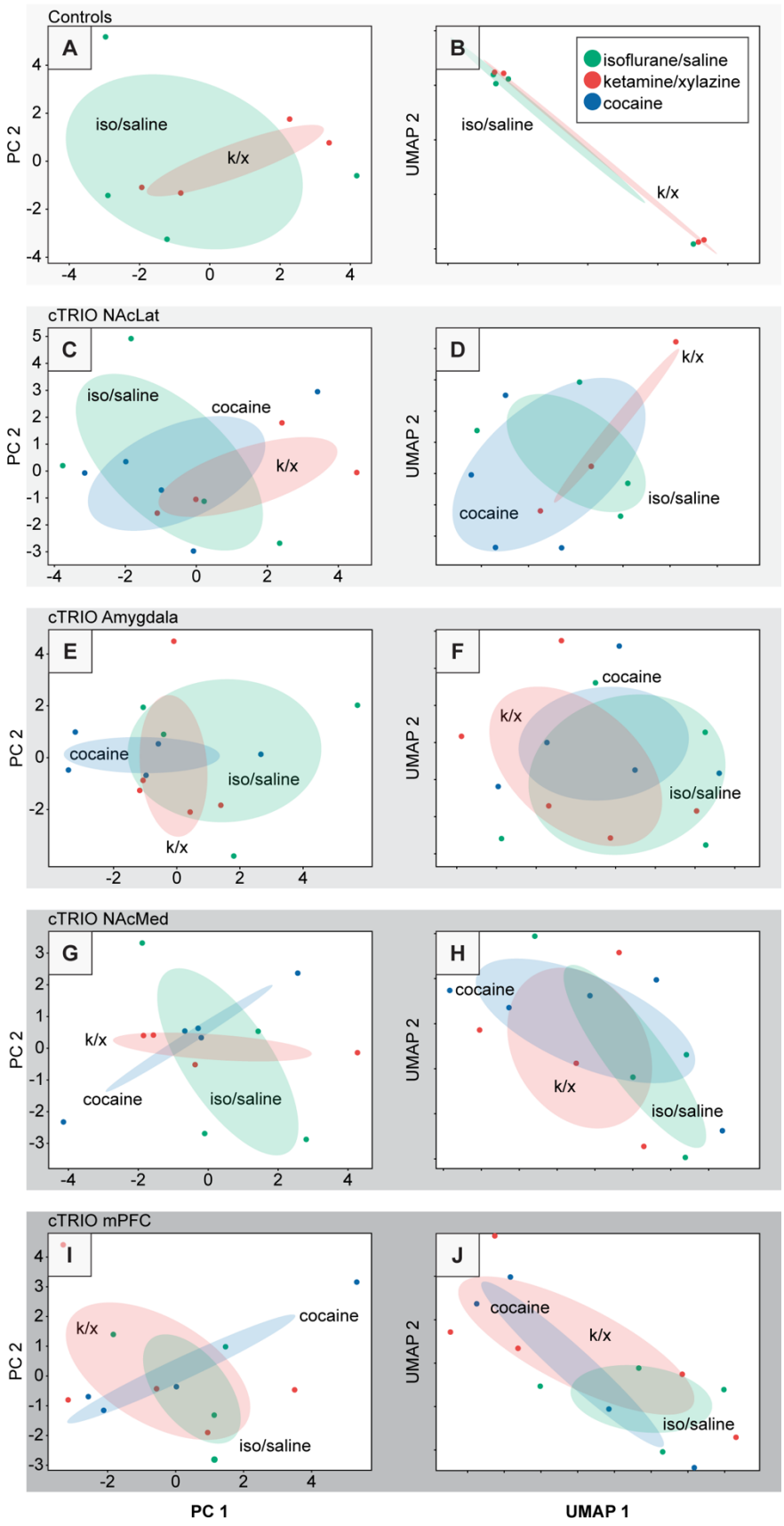


Figure S2.3: Inputs to brains of mice anesthetized with isoflurane and receiving saline or cocaine, or K/X-saline (Figure 2.4), but with the brain identities scrambled.
(A) PCA plot of scrambled data from isoflurane-anesthetized and K/X-anesthetized mice.
(B) UMAP plot of scrambled data from isoflurane-anesthetized and K/X-anesthetized mice.
(C) PCA plot of scrambled data from input mapping experiments to VTA^{DA}→NAcLat cells.
(D) UMAP plot of scrambled data from input mapping experiments to VTA^{DA}→NAcLat cells.
(E) PCA plot of scrambled data from input mapping experiments to VTA^{DA}→Amygdala cells.
(F) UMAP plot of scrambled data from input mapping experiments to VTA^{DA}→Amygdala cells.
(G) PCA plot of scrambled data from input mapping experiments to VTA^{DA}→NAcMed cells.
(H) UMAP plot of scrambled data from input mapping experiments to VTA^{DA}→NAcMed cells.
(I) PCA plot of scrambled data from input mapping experiments to VTA^{DA}→mPFC cells.

(J) UMAP plot of scrambled data from input mapping experiments to $VTA^{DA}\rightarrow$mPFC cells.

Figure S2.4: Gene ontology analysis of top 50 expressed genes correlated to differences in RABV input labeling obtained from drug-treated vs. control mice. GO from mice treated with (A) amphetamine, (B) cocaine, (C) methamphetamine, (D) morphine, and (E) nicotine-treated relative to saline-injected controls are shown. Each GO analysis reflects molecular function-related gene classes whose expression was negatively correlated with RABV labeling differences.

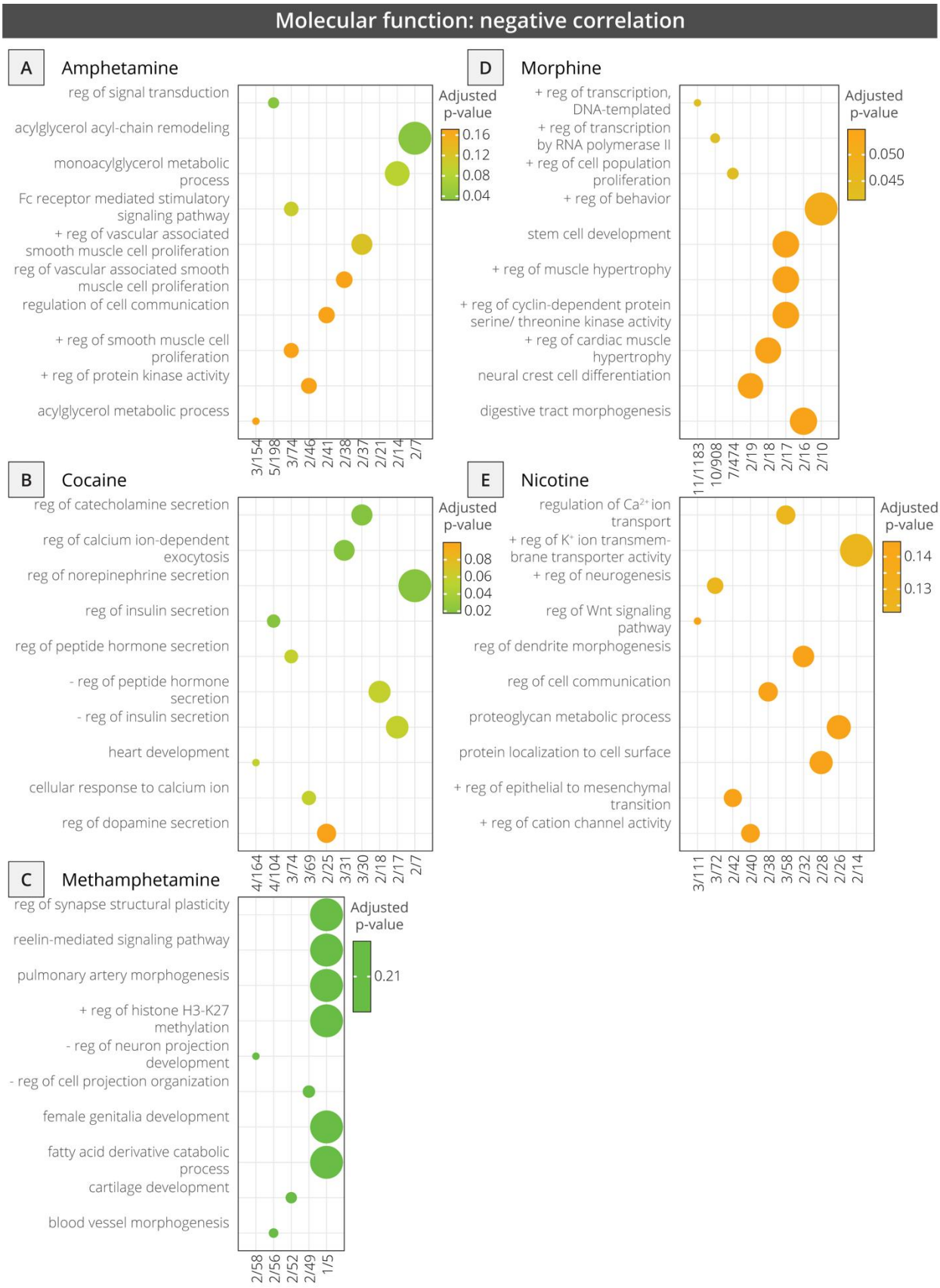
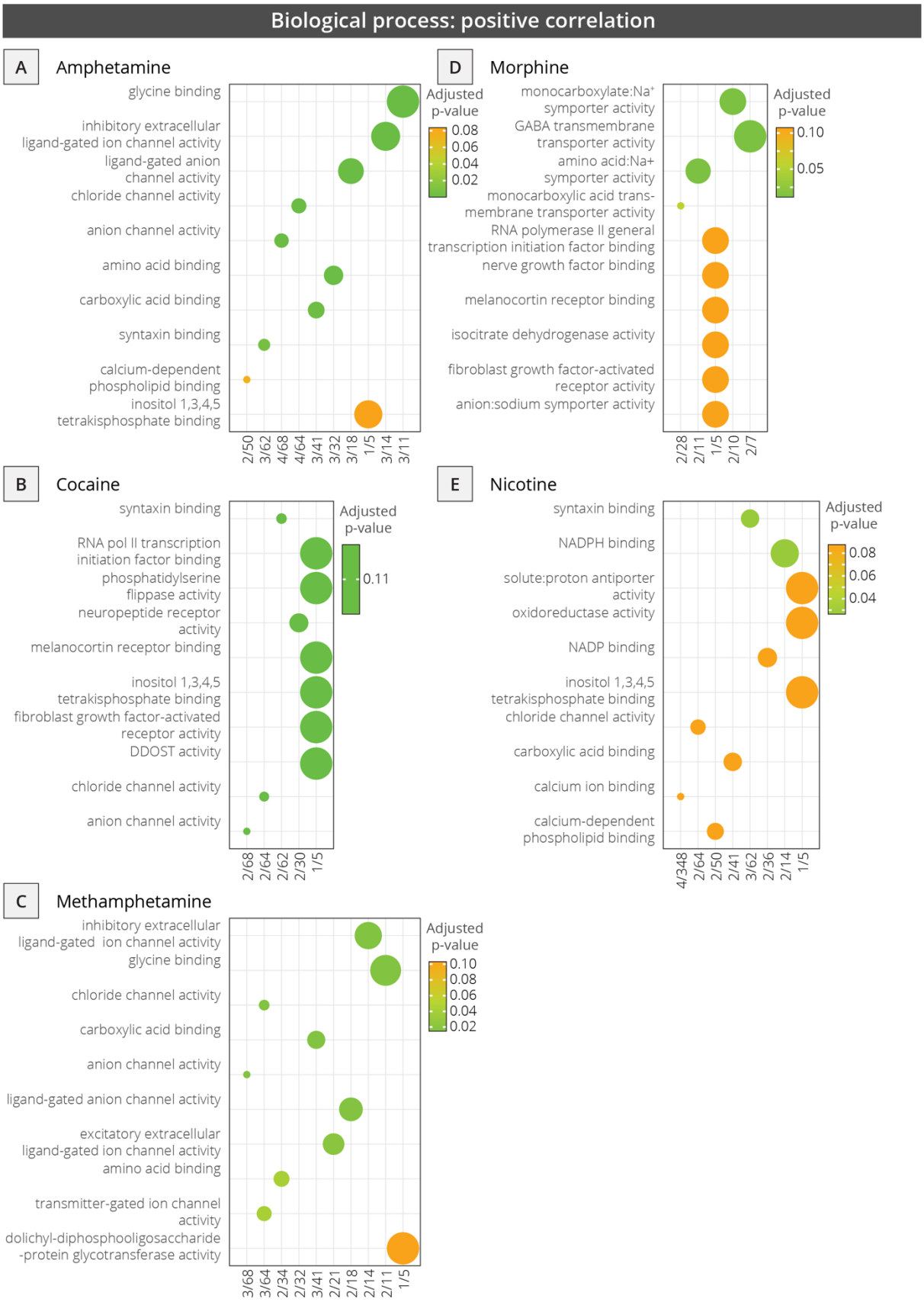

Molecular function: negative correlation
A
Amphetamine
reg of signal transduction
acylglycerol acyl-chain remodeling
monoacylglycerol metabolic process
Fc receptor mediated stimulatory signaling pathway
+ reg of vascular associated smooth muscle cell proliferation
reg of vascular associated smooth muscle cell proliferation
regulation of cell communication
+ reg of smooth muscle cell proliferation
+ reg of protein kinase activity
acylglycerol metabolic process
3/154
5/198
3/74
2/46
2/41
2/38
2/37
2/21
2/14
2/7
Adjusted p-value
0.16
0.12
0.08
0.04
B
Cocaine
reg of catecholamine secretion
reg of calcium ion-dependent exocytosis
reg of norepinephrine secretion
reg of insulin secretion
reg of peptide hormone secretion
- reg of peptide hormone secretion
- reg of insulin secretion
heart development
cellular response to calcium ion
reg of dopamine secretion
4/164
4/104
3/74
3/69
2/25
3/31
3/30
2/18
2/17
2/7
Adjusted p-value
0.08
0.06
0.04
0.02
C
Methamphetamine
reg of synapse structural plasticity
reelin-mediated signaling pathway
pulmonary artery morphogenesis
+ reg of histone H3-K27 methylation
- reg of neuron projection development
- reg of cell projection organization
female genitalia development
fatty acid derivative catabolic process
cartilage development
blood vessel morphogenesis
2/58
2/56
2/52
2/49
1/5
Adjusted p-value
0.21
D
Morphine
+ reg of transcription, DNA-templated
+ reg of transcription by RNA polymerase II
+ reg of cell population proliferation
+ reg of behavior
stem cell development
+ reg of muscle hypertrophy
+ reg of cyclin-dependent protein serine/ threonine kinase activity
+ reg of cardiac muscle hypertrophy
neural crest cell differentiation
digestive tract morphogenesis
11/1183
10/908
7/474
2/19
2/18
2/17
2/16
2/10
Adjusted p-value
0.050
0.045
E
Nicotine
regulation of Ca2+ ion transport
+ reg of K+ ion transmembrane transporter activity
+ reg of neurogenesis
reg of Wnt signaling pathway
reg of dendrite morphogenesis
reg of cell communication
proteoglycan metabolic process
protein localization to cell surface
+ reg of epithelial to mesenchymal transition
+ reg of cation channel activity
3/111
3/72
2/42
2/40
2/38
3/58
2/32
2/28
2/26
2/14
Adjusted p-value
0.14
0.13

Figure S2.5: Gene ontology analysis of top 50 expressed genes correlated to differences in RABV input labeling obtained from drug-treated vs. control mice. GO analyses from mice treated with (A) amphetamine, (B) cocaine, (C) methamphetamine, (D) morphine, and (E) nicotine-treated relative to saline-injected controls are shown. Each GO analysis reflects biological process-related gene classes whose expression was positively correlated with RABV labeling differences.


Biological process: positive correlation
A
Amphetamine
glycine binding
inhibitory extracellular ligand-gated ion channel activity
ligand-gated anion channel activity
chloride channel activity
anion channel activity
amino acid binding
carboxylic acid binding
syntaxin binding
calcium-dependent phospholipid binding inositol 1,3,4,5 tetrakisphosphate binding
2/50 3/62 4/68 4/64 3/41 3/32 3/18 1/5 3/14 3/11
Adjusted p-value
0.08
0.06
0.04
0.02
D
Morphine
monocarboxylate:Na+ symporter activity
GABA transmembrane transporter activity
amino acid:Na+ symporter activity
monocarboxylic acid trans-membrane transporter activity
RNA polymerase II general transcription initiation factor binding
nerve growth factor binding
melanocortin receptor binding
isocitrate dehydrogenase activity
fibroblast growth factor-activated receptor activity
anion:sodium symporter activity
2/28 2/11 1/5 2/10 2/7
Adjusted p-value
0.10
0.05
B
Cocaine
syntaxin binding
RNA pol II transcription initiation factor binding
phosphatidylserine flippase activity
neuropeptide receptor activity
melanocortin receptor binding
inositol 1,3,4,5 tetrakisphosphate binding
fibroblast growth factor-activated receptor activity
DDOST activity
chloride channel activity
anion channel activity
2/68 2/64 2/62 2/30 1/5
Adjusted p-value
0.11
E
Nicotine
syntaxin binding
NADPH binding
solute:proton antiporter activity
oxidoreductase activity
NADP binding
inositol 1,3,4,5 tetrakisphosphate binding
chloride channel activity
carboxylic acid binding
calcium ion binding
calcium-dependent phospholipid binding
4/348 2/64 2/50 2/41 3/62 2/36 2/14 1/5
Adjusted p-value
0.08
0.06
0.04
C
Methamphetamine
inhibitory extracellular ligand-gated ion channel activity
glycine binding
chloride channel activity
carboxylic acid binding
anion channel activity
ligand-gated anion channel activity
excitatory extracellular ligand-gated ion channel activity
amino acid binding
transmitter-gated ion channel activity
dolichyl-diphosphooligosaccharide -protein glycotransferase activity
3/68 3/64 2/34 2/32 3/41 2/21 2/18 2/14 2/11 1/5
Adjusted p-value
0.10
0.08
0.06
0.04
0.02

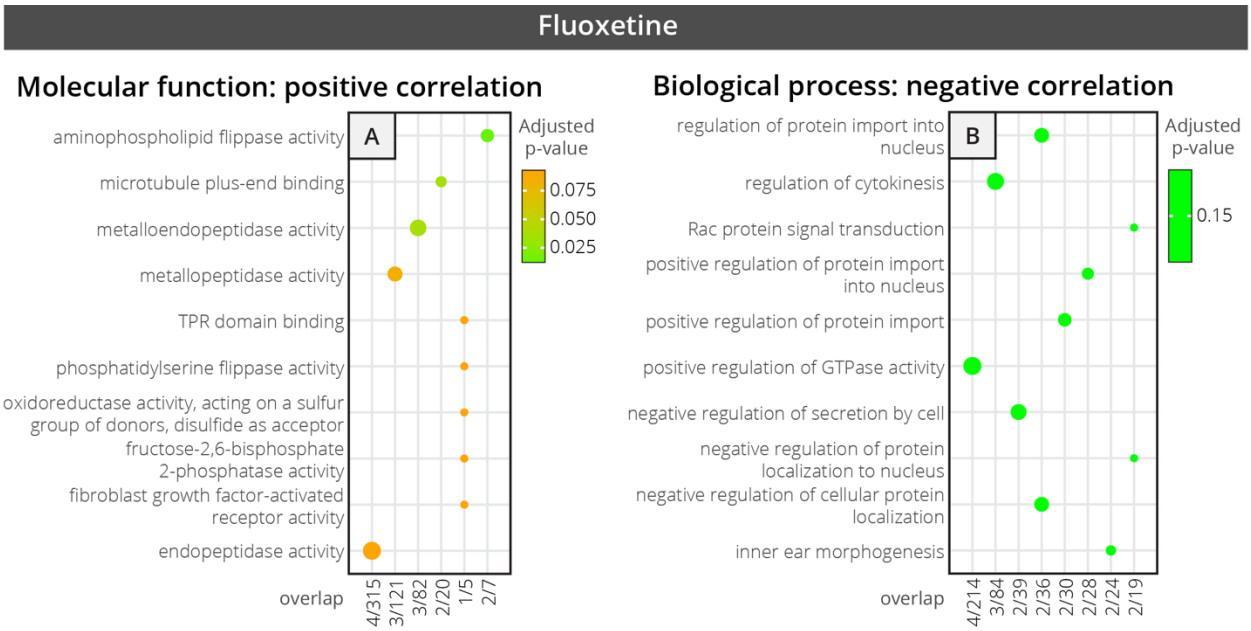


Figure S2.6: Gene ontology analysis of the top 50 genes correlated with differences in RABV input labeling obtained from fluoxetine vs. saline-treated controls. Panel (A) shows molecular function-related gene classes whose expression was positively correlated with RABV labeling differences, and panel (B) shows GO analyses that reflect biological process-related gene classes whose expression was negatively correlated with RABV labeling differences.

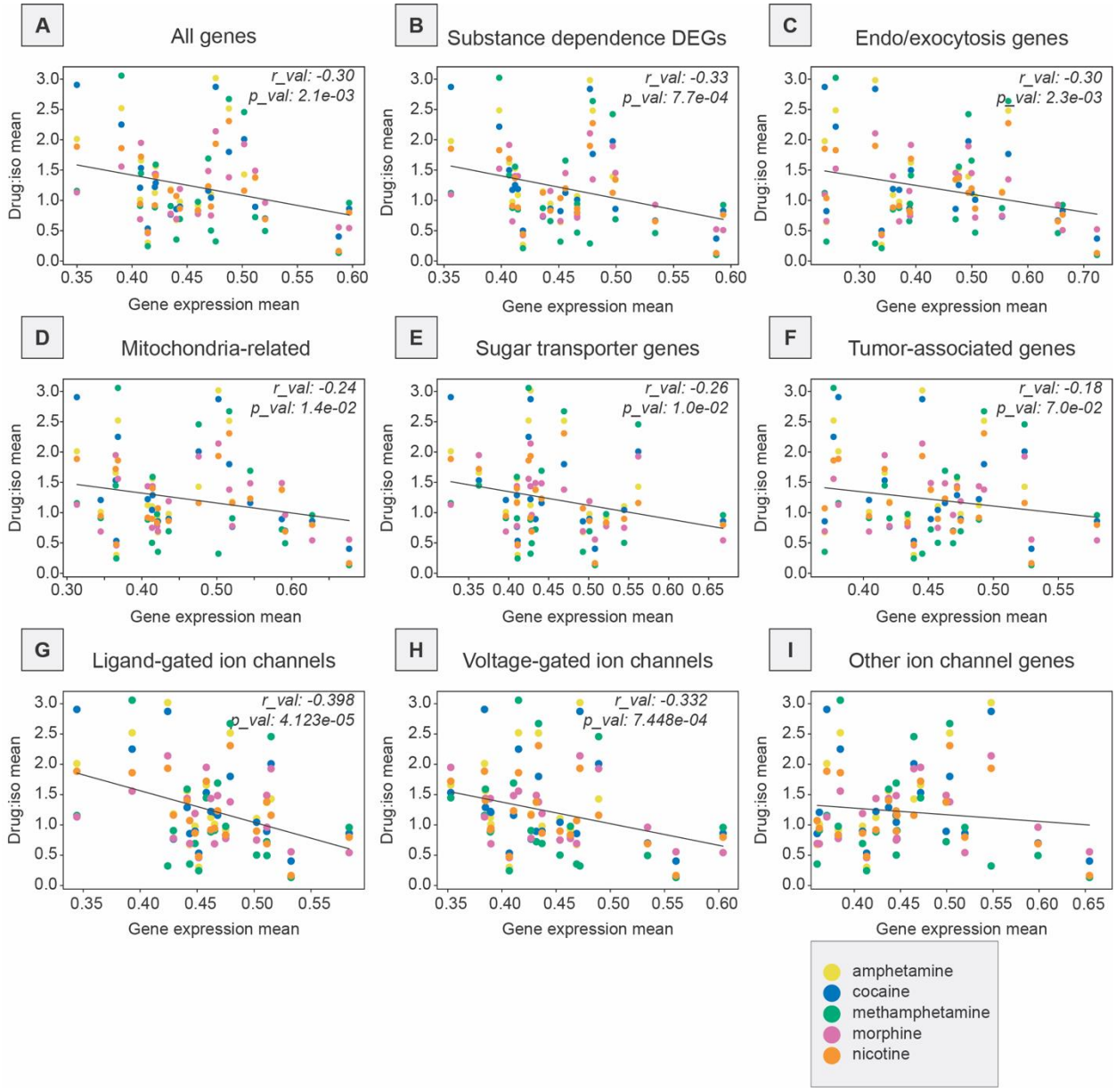


Figure S2.7: Relationship between RABV labeling ratio in addictive drug/saline-treated groups and mean basal gene expression of gene subgroups. Linear regressions are shown for all five examined drugs against all genes and different gene classes expressed by brain region within the Allen Gene Expression Atlas, including (A) all genes, (B) substance dependence-related genes, (C) endocytosis- and exocytosis-related genes, (D) mitochondrial genes, (E) sugar transporter genes, (F) tumor-associated genes, (G) ligand-gated ion channel genes, (H) voltage-gated ion channel genes, (I) other ion channel genes.

Figure S2.8: Relationship between RABV labeling ratio in K/X-saline vs. isoflurane-saline groups and mean basal gene expression of gene subgroups. Linear regressions are shown for K/X-saline vs. isoflurane-saline groups against different gene classes, including (A) ion channel-related genes, (B) synapse-related genes, (C) neurotransmitter-related genes, (D) DEGs related to substance abuse, (E) voltage-gated ion channels, (F) ligand-gated ion channels, (G) other ion channels, (H) K^+ channels, (I) Ca^{2+} channels, (J) Cl^- channels, (K) Na^+ channels, (L) glutamate receptors, (M) glycine receptors, (N) acetylcholine receptors, (O) GABA receptor genes.

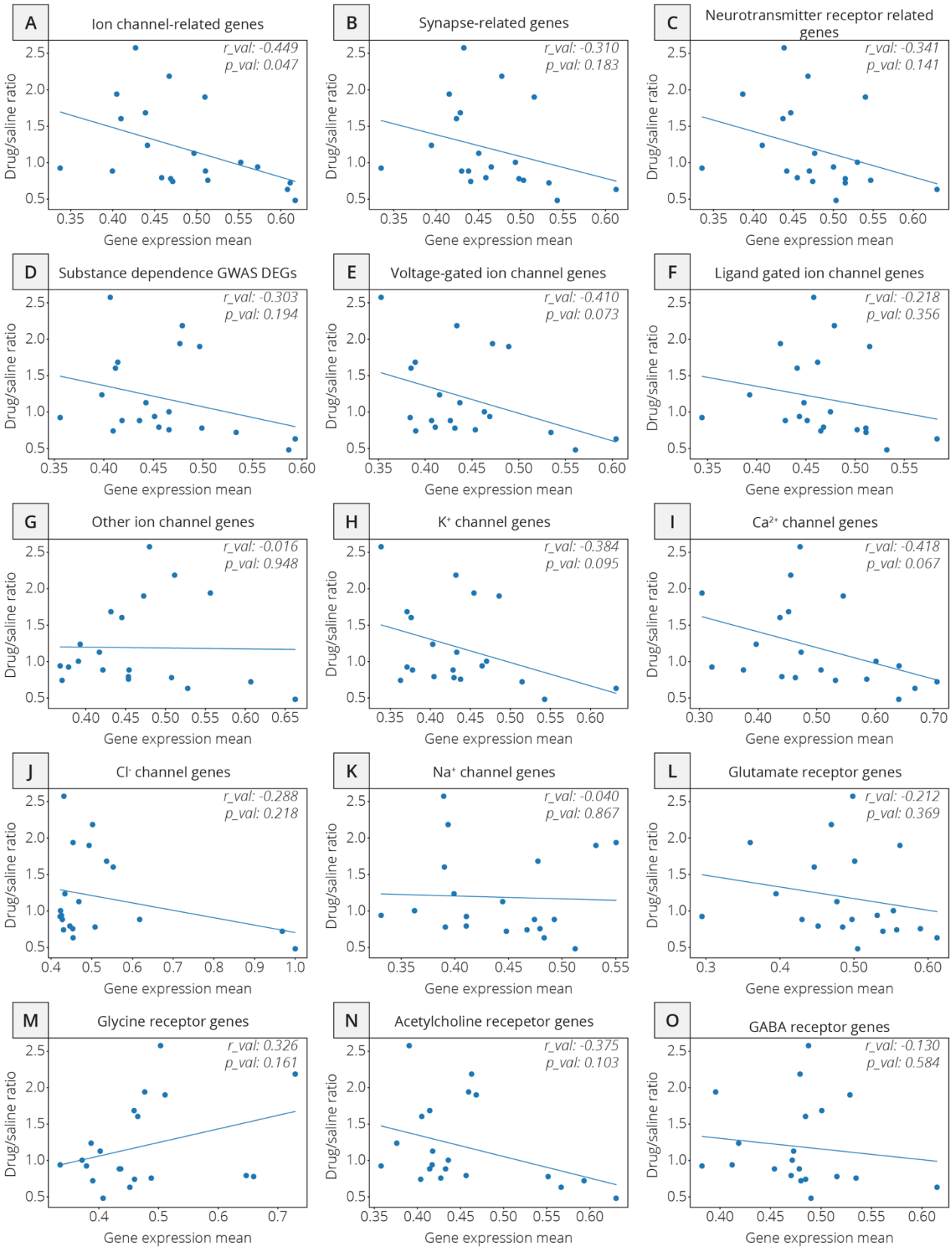
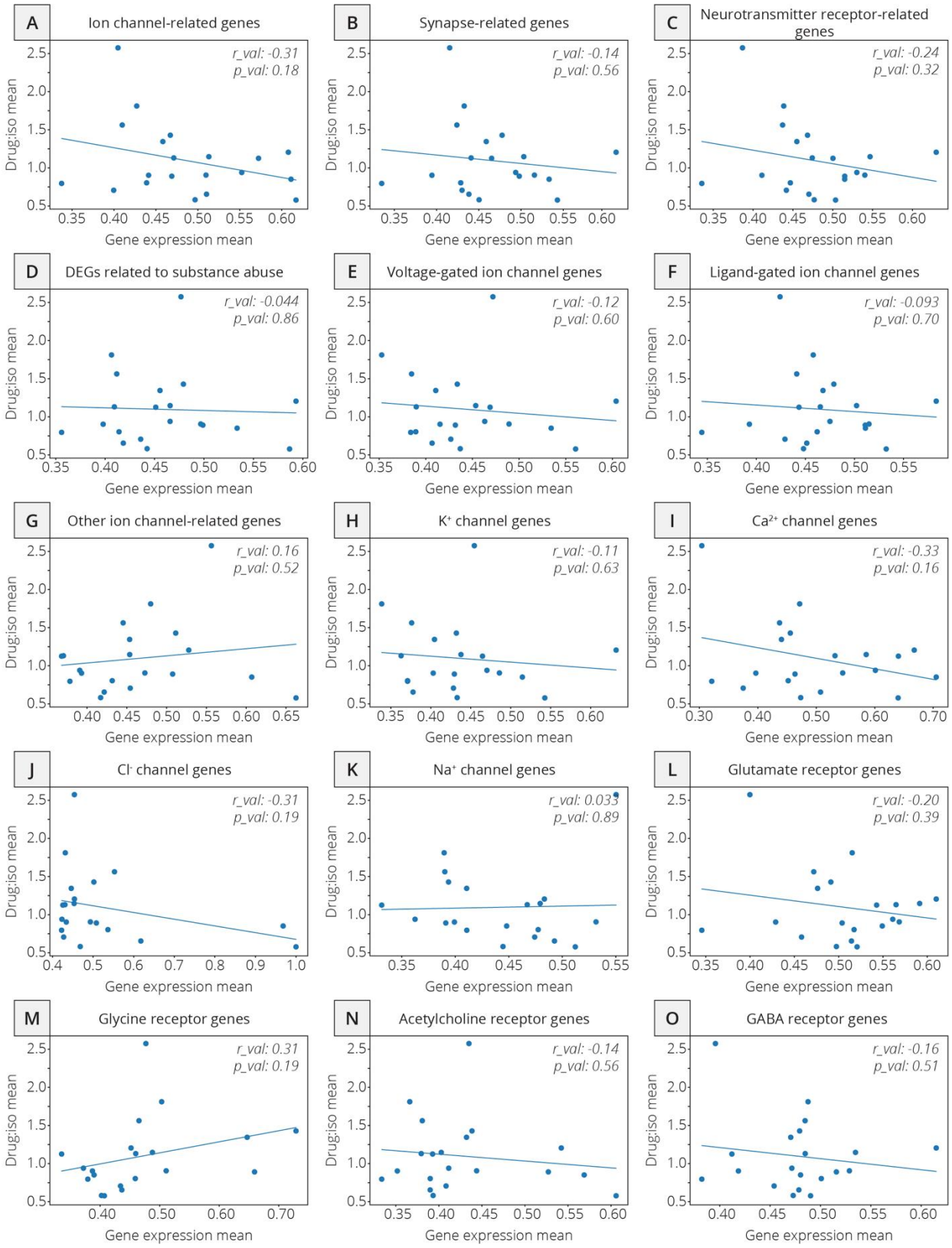

A
Ion channel-related genes
r_val: -0.449
p_val: 0.047
B
Synapse-related genes
r_val: -0.310
p_val: 0.183
C
Neurotransmitter receptor related genes
r_val: -0.341
p_val: 0.141
D
Substance dependence GWAS DEGs
r_val: -0.303
p_val: 0.194
E
Voltage-gated ion channel genes
r_val: -0.410
p_val: 0.073
F
Ligand gated ion channel genes
r_val: -0.218
p_val: 0.356
G
Other ion channel genes
r_val: -0.016
p_val: 0.948
H
K⁺ channel genes
r_val: -0.384
p_val: 0.095
I
Ca²⁺ channel genes
r_val: -0.418
p_val: 0.067
J
Cl⁻ channel genes
r_val: -0.288
p_val: 0.218
K
Na⁺ channel genes
r_val: -0.040
p_val: 0.867
L
Glutamate receptor genes
r_val: -0.212
p_val: 0.369
M
Glycine receptor genes
r_val: 0.326
p_val: 0.161
N
Acetylcholine recepetor genes
r_val: -0.375
p_val: 0.103
O
GABA receptor genes
r_val: -0.130
p_val: 0.584
Drug/saline ratio
Gene expression mean

Figure S2.9: Relationship between RABV ratio labeling in fluoxetine vs. saline-treated groups, both anesthetized with isoflurane, and mean basal gene expression of gene subgroups. Linear regressions are shown for fluoxetine/saline-treated groups against different gene classes, including (A) ion channel-related genes, (B) synapse-related genes, (C) neurotransmitter-related genes, (D) DEGs related to substance abuse, (E) voltage-gated ion channels, (F) ligand-gated ion channels, (G) other ion channels, (H) K^+ channels, (I) Ca^{2+} channels, (J) Cl^- channels, (K) Na^+ channels, (L) glutamate receptors, (M) glycine receptors, (N) acetylcholine receptors, (O) GABA receptor genes.


A
Ion channel-related genes
r_val: -0.31
p_val: 0.18
B
Synapse-related genes
r_val: -0.14
p_val: 0.56
C
Neurotransmitter receptor-related genes
r_val: -0.24
p_val: 0.32
D
DEGs related to substance abuse
r_val: -0.044
p_val: 0.86
E
Voltage-gated ion channel genes
r_val: -0.12
p_val: 0.60
F
Ligand-gated ion channel genes
r_val: -0.093
p_val: 0.70
G
Other ion channel-related genes
r_val: 0.16
p_val: 0.52
H
K⁺ channel genes
r_val: -0.11
p_val: 0.63
I
Ca²⁺ channel genes
r_val: -0.33
p_val: 0.16
J
Cl⁻ channel genes
r_val: -0.31
p_val: 0.19
K
Na⁺ channel genes
r_val: 0.033
p_val: 0.89
L
Glutamate receptor genes
r_val: -0.20
p_val: 0.39
M
Glycine receptor genes
r_val: 0.31
p_val: 0.19
N
Acetylcholine receptor genes
r_val: -0.14
p_val: 0.56
O
GABA receptor genes
r_val: -0.16
p_val: 0.51
Drug:iso mean
Gene expression mean

2.6 References

1. Bromberg-Martin, E. S., Matsumoto, M. & Hikosaka, O. Dopamine in Motivational Control: Rewarding, Aversive, and Alerting. *Neuron* **68**, 815–834 (2010).
2. Lammel, S., Lim, B. K. & Malenka, R. C. Reward and aversion in a heterogeneous midbrain dopamine system. *Neuropharmacology* **76**, 351–359 (2014).
3. Wise, R. A. Dopamine, learning and motivation. *Nat Rev Neurosci* **5**, 483–494 (2004).
4. Kim, C. K. *et al.* Simultaneous fast measurement of circuit dynamics at multiple sites across the mammalian brain. *Nat Methods* **13**, 325–8 (2016).
5. Lammel, S., Ion, D. I., Roeper, J. & Malenka, R. C. Report Projection-Specific Modulation of Dopamine Neuron Synapses by Aversive and Rewarding Stimuli. *Neuron* **70**, 855–862 (2011).
6. Lammel, S. *et al.* Input-specific control of reward and aversion in the ventral tegmental area. *Nature* **491**, 212–217 (2012).
7. Beier, K. T. *et al.* Circuit Architecture of VTA Dopamine Neurons Revealed by Systematic Input-Output Mapping. *Cell* **162**, 622–634 (2015).
8. Menegas, W. *et al.* Dopamine neurons projecting to the posterior striatum form an anatomically distinct subclass. *Elife* **4**, 1–30 (2015).
9. Watabe-Uchida, M., Zhu, L., Ogawa, S. K., Vamanrao, A. & Uchida, N. Whole-Brain Mapping of Direct Inputs to Midbrain Dopamine Neurons. *Neuron* **74**, 858–873 (2012).
10. Beier, K. T. *et al.* Topological Organization of Ventral Tegmental Area Connectivity Revealed by Viral-Genetic Dissection of Input-Output Relations. *Cell Rep* **26**, 159-167.e6 (2019).
11. Cardozo Pinto, D. F. *et al.* Characterization of transgenic mouse models targeting neuromodulatory systems reveals organizational principles of the dorsal raphe. *Nat Commun* **10**, (2019).
12. Faget, L. *et al.* Afferent Inputs to Neurotransmitter-Defined Cell Types in the Ventral Tegmental Area. *Cell Rep* **15**, 2796–2808 (2016).
13. Kalivas, P. W. & Volkow, N. D. The neural basis of addiction: a pathology of motivation and choice. *Am J Psychiatry* **162**, 1403–1413 (2005).
14. Hyman, S. E., Malenka, R. C. & Nestler, E. J. Neural mechanisms of addiction: the role of reward-related learning and memory. *Annu Rev Neurosci* **29**, 565–598 (2006).
15. Beier, K. T. *et al.* Rabies screen reveals GPe control of cocaine-triggered plasticity. *Nature* **549**, 345–350 (2017).
16. Tian, G. *et al.* An extended Amygdal-midbrain circuit controlling cocaine withdrawal-induced anxiety and reinstatement. *Cell Rep* **39**, 1–16 (2022).
17. Juarez, B. & Han, M. H. Diversity of Dopaminergic Neural Circuits in Response to Drug Exposure. *Neuropsychopharmacology* **41**, 2424–2446 (2016).
18. Mahler, S. V *et al.* Chemogenetic Manipulations of Ventral Tegmental Area Dopamine Neurons Reveal Multifaceted Roles in Cocaine Abuse. *Journal of Neuroscience* **39**, 503–518 (2019).
19. Saal, D., Dong, Y., Bonci, A. & Malenka, R. C. Drugs of abuse and stress trigger a common synaptic adaptation in dopamine neurons. *Neuron* **37**, 577–582 (2003).
20. Ungless, M. A., Whistler, J. L., Malenka, R. C. & Bonci, A. Single cocaine exposure in vivo induces long-term potentiation in dopamine neurons. *Nature* **411**, 583–7. (2001).
21. Niehaus, J. L., Murali, M. & Kauer, J. A. Drugs of abuse and stress impair LTP at inhibitory synapses in the ventral tegmental area. *European Journal of Neuroscience* **32**, 108–117 (2010).

22. Bocklisch, C. *et al*. Cocaine Disinhibits Dopamine Neurons by Potentiation of GABA Transmission in the Ventral Tegmental Area. *Science (1979)* **341**, 1521–1525 (2013).
23. Dong, Y. *et al*. Cocaine-induced potentiation of synaptic strength in dopamine neurons: behavioral correlates in GluRA(-/-) mice. *Proc Natl Acad Sci U S A* **101**, 14282–14287 (2004).
24. Creed, M. *et al*. Cocaine exposure enhances the activity of ventral tegmental area dopamine neurons via calcium-impermeable NMDARs. *Journal of Neuroscience* **36**, 10759–10768 (2016).
25. Tian, G. *et al*. An extended Amygdala-midbrain circuit controlling cocaine withdrawal-induced anxiety and reinstatement. *Cell Rep* **39**, 1–16 (2022).
26. Backman, C. M. *et al*. Characterization of a mouse strain expressing Cre recombinase from the 3' untranslated region of the dopamine transporter locus. *Genesis* **44**, 383–390 (2006).
27. Taniguchi, H. *et al*. A Resource of Cre Driver Lines for Genetic Targeting of GABAergic Neurons in Cerebral Cortex. *Neuron* **71**, 995–1013 (2011).
28. Vong, L. *et al*. Leptin Action on GABAergic Neurons Prevents Obesity and Reduces Inhibitory Tone to POMC Neurons. *Neuron* **71**, 142–154 (2011).
29. Kumar, N. *et al*. The development of an AAV-based crispr sacas9 genome editing system that can be delivered to neurons in vivo and regulated via doxycycline and cre-recombinase. *Front Mol Neurosci* **11**, 1–14 (2018).
30. Pedregosa, F. *et al*. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **12**, 2825–2830 (2011).
31. McInnes, L., Healy, J. & Melville, J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *ArXiv* **1802.03426**, 1–51 (2018).
32. Derdeyn, P., Hui, M., Macchia, D. & Beier, K. Uncovering the Connectivity Logic of the Ventral Tegmental Area. *Front Neural Circuits* **15**, 799688 (2022).
33. Waskom, M. Seaborn: Statistical Data Visualization. *J Open Source Softw* **6**, 3021 (2021).
34. Hunter, J. D. Matplotlib: a 2D graphics environment. *Comput Sci Eng* **9**, 90–95 (2007).
35. Chon, U., Vanselow, D. J., Cheng, K. C. & Kim, Y. Enhanced and unified anatomical labeling for a common mouse brain atlas. *Nat Commun* **10**, (2019).
36. Oh, S. W. *et al*. A mesoscale connectome of the mouse brain. *Nature* **508**, 207–214 (2014).
37. Lein, E. S. *et al*. Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445**, 168–176 (2007).
38. Brynildsen, J. K. *et al*. Gene coexpression patterns predict opiate-induced brain-state transitions. *Proc Natl Acad Sci U S A* **117**, 19556–19565 (2020).
39. Fulcher, B. D., Murray, J. D., Zerbi, V. & Wang, X. J. Multimodal gradients across mouse cortex. *Proc Natl Acad Sci U S A* **116**, 4689–4695 (2019).
40. Sian, J., Youdim, M. & Riederer, P, et al. Biochemical Anatomy of the Basal Ganglia and Associated Neural Systems. in *Basic Neurochemistry: Molecular, Cellular and Medical Aspects.* (eds. Siegel, G., Agranoff, B. & Albers, RW, et al.) (1999).
41. Xie, Z. *et al*. Gene Set Knowledge Discovery with Enrichr. *Curr Protoc* **1**, 1–51 (2021).
42. Ji, X., Zhang, H. & Cui, Q. A Panel of Synapse-Related Genes as a Biomarker for Gliomas. *Front Neurosci* **14**, 1–8 (2020).
43. Wang, R. *et al*. Ion channel gene expression predicts survival in glioma patients. *Sci Rep* **5**, 1–10 (2015).
44. Berchtold, N. C. *et al*. Synaptic genes are extensively downregulated across multiple brain regions in normal human aging and Alzheimer's disease. *Neurobiol Aging* **34**, 1653–1661 (2013).
45. Shi, L. *et al*. AddictGene: An integrated knowledge base for differentially expressed genes associated with addictive substance. *Comput Struct Biotechnol J* **19**, 2416–2422 (2021).

46. Saheki, Y. & De Camilli, P. Synaptic vesicle endocytosis. *Cold Spring Harb Perspect Biol* **4**, (2012).
47. Harding, S. D. *et al.* The IUPHAR/BPS guide to PHARMACOLOGY in 2022: Curating pharmacology for COVID-19, malaria and antibacterials. *Nucleic Acids Res* **50**, D1282–D1294 (2022).
48. Yamaguchi, T., Wang, H. L., Li, X., Ng, T. H. & Morales, M. Mesocorticolimbic glutamatergic pathway. *Journal of Neuroscience* **31**, 8476–8490 (2011).
49. Yamaguchi, T., Sheen, W. & Morales, M. Glutamatergic neurons are present in the rat ventral tegmental area. *European Journal of Neuroscience* **25**, 106–118 (2007).
50. Root, D. H. *et al.* Single rodent mesohabenular axons release glutamate and GABA. *Nat Neurosci* **17**, 1543–1551 (2014).
51. Morales, M. & Margolis, E. B. Ventral tegmental area: Cellular heterogeneity, connectivity and behaviour. *Nat Rev Neurosci* **18**, 73–85 (2017).
52. Hnasko, T. S., Hjelmstad, G. O., Fields, H. L. & Edwards, R. H. Ventral tegmental area glutamate neurons: electrophysiological properties and projections. *Journal of Neuroscience* **32**, 15076–15085 (2012).
53. Kim, J. I. *et al.* Aldehyde dehydrogenase 1a1 mediates a GABA synthesis pathway in midbrain dopaminergic neurons. *Science (1979)* **350**, 102–106 (2015).
54. Tritsch, N. X., Ding, J. B. & Sabatini, B. L. Dopaminergic neurons inhibit striatal output through non-canonical release of GABA. *Nature* **490**, 262–266 (2012).
55. Koob, G. F. A Role for Brain Stress Systems in Addiction. *Neuron* **59**, 11–34 (2008).
56. Koob, G. F. Brain stress systems in the amygdala and addiction. *Brain Res* **1293**, 61–75 (2009).
57. Lammel, S. *et al.* Unique Properties of Mesoprefrontal Neurons within a Dual Mesocorticolimbic Dopamine System. *Neuron* **57**, 760–773 (2008).
58. Larson, M. H. *et al.* CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nat Protoc* **8**, 2180–2196 (2013).
59. Li, H. *et al.* Entopeduncular nucleus projections to the lateral habenula contribute to cocaine avoidance. *Journal of Neuroscience* **41**, 298–306 (2021).
60. López, A. J. *et al.* Medial habenula cholinergic signaling regulates cocaine-associated relapse-like behavior. *Addiction Biology* **24**, 403–413 (2019).
61. Tan, D., Nuno-Perez, A., Mameli, M. & Meye, F. J. Cocaine withdrawal reduces GABABR transmission at entopeduncular nucleus – lateral habenula synapses. *European Journal of Neuroscience* **50**, 2124–2133 (2019).
62. Koob, G. F. & Volkow, N. D. Neurocircuitry of addiction. *Neuropsychopharmacology* **35**, 217–238 (2010).
63. Tian G *et al.* Molecular and circuit determinants in the globus pallidus mediating control of cocaine-induced behavioral plasticity. *Neuron Sneak Peek* (2023).
64. Liu, Y., Lin, D., Wu, B. & Zhou, W. Ketamine abuse potential and use disorder. *Brain Res Bull* **126**, 68–73 (2016).
65. Ni, A., Lee Cantrell, F. & Clark, R. F. Ketamine exposure demographics and outcomes over 16 years as reported to US poison centers. *American Journal of Emergency Medicine* **36**, 1459–1462 (2018).
66. Dillon, P., Copeland, J. & Jansen, K. Patterns of use and harms associated with non-medical ketamine use. *Drug Alcohol Depend* **69**, 23–28 (2003).
67. Morgan, C. J. A. & Curran, H. V. Ketamine use: A review. *Addiction* **107**, 27–38 (2012).

68. Ryder, S., Way, W. L. & Trevor, A. J. Comparative pharmacology of the optical isomers of ketamine in mice. *Eur J Pharmacol* **49**, 15–23 (1978).
69. Finck, A. D. & Ngai, S. H. Opiate receptor mediation of ketamine analgesia. *Anesth Analg* **56**, 291–297 (1982).
70. Hustveit, O., Maurset, A. & Oye, I. Interaction of the Chiral Forms of Ketamine with Opioid, Phencyclidine, c and Muscarinic Receptors. *Pharmacol Toxicol* **77**, 355–359 (1995).
71. Bonaventura, J. *et al.* Pharmacological and behavioral divergence of ketamine enantiomers:implications for abuse liability. *Mol Psychiatry* **26**, 6704–6722 (2021).
72. Williams, N. R. *et al.* Attenuation of antidepressant and antisuicidal effects of ketamine by opioid receptor antagonism. *Mol Psychiatry* **24**, 1779–1786 (2019).
73. Williams, N. R. *et al.* Attenuation of antidepressant effects of ketamine by opioid receptor antagonism. *American Journal of Psychiatry* **175**, 1205–1215 (2018).
74. Gelernter, J. *et al.* Genome-wide association study of opioid dependence: Multiple associations mapped to calcium and potassium pathways. *Biol Psychiatry* **76**, 66–74 (2014).
75. Malcolm, R., Brady, K. T., Moore, J. & Kajdasz, D. Amlodipine treatment of cocaine dependence. *J Psychoactive Drugs* **31**, 117–120 (1999).
76. Altamura, A. C., Regazzetti, M. G. & Porta, M. Nimodipine in human alcohol withdrawal syndrome - an open study. *European Neuropsychopharmacology* **1**, 37–40 (1990).
77. Rosse, R. *et al.* Nimodipine Pharmacotherapeutic Adjuvant Therapy for Inpatient Treatment of Cocaine Dependence. *Clin Neuropharmacol* **17**, 348–358 (1994).
78. Muntaner, C., Kumor, K., Nagoshi, C. & Jaffe, J. Effect of nifedipine pretreatment on subjective and cardiovascular responses to intravenous cocaine in humans. *Drug Alcohol Depend* **105**, 37–41 (1991).
79. Johnson, B. A. *et al.* Effects of Isradipine on Cocaine-Induced Subjective Mood. *J Clin Psychopharmacol* **24**, 180–191 (2004).
80. Vaupel, D. B., Lange, W. R. & London, E. D. Effects of verapamil on morphine-induced euphoria, analgesia and respiratory depression in humans. *Journal of Pharmacology and Experimental Therapeutics* **267**, 1386–1394 (1993).

CHAPTER 3: Differences in cellular communication in the entorhinal cortex in AD and in 5xFAD mice

As a precursor to studying changes in the retrosplenial cortex, our first analysis of AD and the 5xFAD mouse model focused on the entorhinal cortex due to the known early changes that occur in this brain region in AD progression, as discussed in chapter 1 of this dissertation. Additionally, due to the previous study of this brain region, there was publicly available snRNA-seq data from donors with and without AD that we utilized for re-analysis and comparison with 5xFAD data we collected. We detail inferred cellular communication changes that are common between the mouse model and post-mortem tissues samples, and later utilize similar approaches to compare the RSC and the ENT. The findings from this chapter highlight the importance of changes in glutamatergic neurons in AD and in 5xFAD mice, which we delve into further with additional experiments and sequencing data collection in chapter 4.

Chapter 3 of this dissertation is currently in review. I performed the majority of computational analyses of the data, Dr. May Hui and Dr. Desiree Macchia assisted with figure preparation, Dr. Wei Zhao and Dr. Qing Nie reviewed and advised on computational approaches, Dr. Kevin T. Beier performed sequencing experiments, and Dr. Kevin T. Beier and I wrote the manuscript.

3.1 Introduction

Alzheimer's disease (AD) is a complex disease characterized by systemic changes in brain function[1,2]. While the pathological hallmarks of AD are the accumulation of amyloid plaques and tau tangles in the brain, changes in cellular function occur in neurons and glia throughout the course of disease that contribute to the development of age-associated cognitive impairment[3,4]. Understanding the root molecular causes of disease is thus required for the development of

effective therapeutics to combat the development of AD. Our ability to understand the molecular features of disease across the lifespan has been greatly accelerated by the recent development of single-cell and single-nucleus RNA sequencing (snRNA-seq) technologies. These methods have provided an atlas of gene expression throughout the course of disease, particularly in late-stage AD in brain regions such as the frontal cortex, entorhinal cortex (ENT), and the hippocampus[5–8]. However, a key limitation of these experiments is that they focus only on the cell-autonomous features of the disease; that is, they examine cells in isolation. While a reductionist approach is necessary to define the core features of disease, understanding how the molecular features of disease, as revealed by snRNA-seq, influence cellular communication is essential towards uncovering the root and progression of AD.

Fortunately, in recent years a variety of computational methods have been developed that enable the inference of cellular communication through coordinated changes in receptor-ligand interactions in different cell types. These include CellPhoneDB[9] and CellChat[10], among others. These methods have been widely used in a variety of tissue types to enrich single cell sequencing datasets and provide a wealth of information about modes of biological communication within tissues including in the AD brain. However, as these studies focused on glia and did not analyze data from neurons, they include only data from selected cell types[11–15]. While these methods' utility in part is due to their ability to be generalized to different tissues, they also focus exclusively on receptor-ligand interactions that are encoded by proteins. While this is an acceptable definition for most tissue types, its use is more limited in the brain, where cells principally communicate through chemical transmission of small-molecule neurotransmitters, including glutamate and GABA, that are not encoded by proteins. To fill this gap, NeuronChat[16] was developed to infer the intercellular communications mediated by small-molecule neurotransmitters. As such, if we want

to understand how communication systems in the brain change during disease, an integrated understanding of how both general communication and neuron-specific modes of communication change during disease is required.

Here, we focus on the ENT, as it is a core region in the hippocampal circuit and thought to be a main contributor to AD-associated memory and cognitive deficits[17–20]. For these studies we used the 5xFAD mouse model, a commonly studied amyloidogenic model used in ~10% of all AD studies employing rodent models[21,22]. We sequenced the ENT from 5xFAD and littermate control mice at 2 months of age, before memory deficits are apparent, and at 8 months of age, when behavioral deficits are widespread[22,23]. For comparison of gene expression changes in rodents to human postmortem brain tissue we utilized publicly available snRNA-seq data from the ENT of patients with AD and those without cognitive impairments. We also sequenced postmortem tissue from the ENT of a patient diagnosed with mild cognitive impairment (MCI) to provide a more comprehensive examination of AD progression in mice and human postmortem tissue across disease progression.

3.2 Methods

Mouse and human tissue

Hemizygous 5xFAD mice (mutations in APP: Swedish, Florida, and London; mutations in PSEN1: M146L and L286V) on a C57BL/6 background (Jackson Laboratory, Stock #34848) were utilized. Wildtype littermates on the same background served as control animals. Female mice were used exclusively for all experiments. Mice were euthanized at either 2 or 8 months of age depending on the time point of the experiment. Mice were housed on a 12-hour light–dark cycle with food and water ad libitum. All procedures complied with the animal care standards set forth by the National Institute of Health and were approved by the University of California, Irvine's

Institutional Animal Care and Use Committee (IACUC). Human postmortem brain tissue samples were obtained from UC Irvine's Alzheimer's Disease Research Center (ADRC). The female donor was diagnosed with mild cognitive impairment (MCI), had tau tangle stage 2, plaque stage A, and was 85 years old (sample 28_18).

Sample collection and single-nucleus sequencing in mice

Single-nucleus ATAC-seq and RNA-seq were conducted using the 10x Multiome kit. A merged sample comprising nuclei from two female 5xFAD mice and another merged sample from two female C57BL/6 wild-type mice, either at 2 months or 8 months of age, were used, for a total of eight mice.

Animals were anesthetized under isoflurane induction and underwent cardiac perfusion with cold 1xPBS to remove blood from the brain. Brain tissue was immediately extracted, and entorhinal cortex was freshly dissected out of each sample. From this point onwards, both mouse and human tissue were treated the same. To isolate the nuclei, tissue was homogenized using Lysis Buffer and manually disrupted using pestles, followed by a 4-minute incubation period. Samples were centrifuged at 500g for 5 minutes at 4°C, and the supernatant was discarded. Subsequently, 1 ml of Nuclei Wash and Resuspension buffer (NWR) was added and incubated for 5 minutes. The pellet was carefully resuspended using a P1000 pipette and centrifuged again at 500g for 5 minutes at 4°C. This wash-resuspension-centrifugation process was repeated twice more. After the third centrifugation, the supernatant was discarded, and 200µL of NWR buffer + 20µL 7AAD (1:100) was added and passed through a 40µm filter directly into a tube used for flow cytometry sorting. Cells were sorted at the UC Irvine Institute for Immunology Flow Cytometry Facility to remove debris and distinguish viable nuclei. Following sorting, the samples underwent a permeabilization process, which began with centrifugation at 500g for 5 minutes at 4°C. Subsequently, 100µL of

0.1X Lysis buffer was added, followed by resuspension and a 2-minute incubation. Next, 1mL of Washing Buffer was added, and the sample was centrifuged again for 5 minutes under the same conditions. Then, 500µL of Diluted Nuclei Buffer was added and carefully mixed with a pipette. Finally, the resulting sample was spun down and submitted to the UC Irvine's Genomics Research and Technology Hub for further processing.

Single-nucleus sequencing alignment and QC in mice

Alignment of the ATAC-seq and gene expression libraries to the mouse reference genome (mm10, v2020-A) was done using Cell Ranger ARC (v2.0.1, 10X Genomics). Alignment was performed on the High-Performance Community Computing Cluster (HPC3) at UC Irvine. Further processing was done in R largely using Signac[24] (v1.7.0), which is integrated with Seurat[25] (v4.3.0). Nuclei with the following QC metrics in the RNA-seq data were filtered out as they were likely doublets, debris, or damaged: over 20 percent of reads aligned to mitochondrial genes, more than 70,000 or fewer than 200 UMIs, or more than 10,000 or fewer than 75 unique genes detected (**Figure S3.1A-B**). Nuclei with the following metrics were further filtered out from the ATAC-seq data, to prioritize only high-quality nuclei: fewer than 10 percent reads in peaks, nucleosome signal score of greater than 3, transcription start site enrichment of less than 2, over 5 percent of reads being in a blacklist region, or counts in peaks of under 1,000 or over 70,000.

Following QC, there were a total of 35,188 high quality nuclei remaining: 8,286 from 2-month-old C57BL/6 mice, 15,126 from 2-month-old 5xFAD mice, 6,431 from 8-month-old C57BL/6 mice, and 5,345 from 8-month-old 5xFAD mice.

Single-nucleus sequencing alignment and QC, human postmortem samples

Previously published snRNA-seq data collected from the ENT of AD and non-cognitively impaired, or control, patients was downloaded from the gene expression omnibus[5,26] and re-

processed using Seurat. Only female patient data was included for consistency with our other dataset from exclusively female mice. From these two previously published datasets, data from 8 patients were used: 4 female patients with AD, and 4 female patients without cognitive impairments. The additional female MCI patient sample was sequenced at UCI's GRT Hub and resulting fastqs aligned to the NCBI Homo Sapiens Annotation Release 109 using the Parse Biosciences computational pipeline (v1.0.5p). The resulting gene by cell counts matrices were QCed individually by sample and subsequently integrated using Seurat's IntegrateData function with default parameters.

Single-nucleus sequencing integration, dimensionality reduction, and clustering, mice

To integrate 5xFAD and control mouse brain transcriptomic data, sctransform-based normalization was used to normalize, scale, and find variable features within each dataset[27]. Mitochondrial mapping percentage, a potential confounder, was regressed out, and the Gamma-Poisson Generalized Linear Model method[28] was used, otherwise default sctransform parameters were used. 3,000 features that varied the most across all datasets were selected as integration features, and the datasets were integrated using the IntegrateData command in Seurat, with anchor features being set to our list of 3,000 features, the normalize method specified as sctransform, and the other parameters set to their default values. This resulted in one Signac object containing all mouse data, including the 2-month-old and 8-month-old 5xFAD and WT mice. After integration, the standard Seurat workflow for dimensionality reduction, visualization, and clustering was run sequentially using the following commands: RunPCA, RunUMAP, FindNeighbors, and FindClusters. For RunUMAP and FindNeighbors the first 30 principal components were used, and for RunUMAP the number of nearest neighbors parameter was set to 30. The resolution parameter for FindClusters was set to 1.0.

Annotation of cell types

Cell types in our mouse data were predicted by transferring annotation labels from the Allen Brain Atlas' transcriptomic data[29,30] to our datasets, and this predicted annotation was manually checked and refined using previously established marker genes[31–33]. Oligodendrocyte markers used were *Mobp*, *Il33* and *Mog*, OPC markers were *Cspg4*, *Tnr*, and *Pdgfra*, astrocyte markers were *Aqp4*, *Slc1a2*, and *Gja1*, microglia markers were *Csf1r*, *Ctss*, and *C1qb*, GABAergic neuron markers were *Gad1* and *Gad2*, glutamatergic neuron markers were *Slc17a7* and *Slc17a6*, and endothelial cell markers were *Bsg*, *Flt1*, and *Vwf*. The same marker genes were used to manually annotate the data from postmortem human samples. Clusters of fewer than 100 nuclei, that did not express known marker genes, or that expressed marker genes of multiple cells types, indicating a mixed population, were removed.

snATAC-seq data processing, mice

Following integration and annotation, we re-processed the ATAC data. As peaks had previously been called on each sample individually by Cell Ranger ARC, they needed to be redefined so that common peaks would be identified across all mouse brain samples. We also aimed to call peaks by cell type to identify peaks in rare cell types that were potentially missed in aggregate. On the integrated data we ran the CallPeaks function in Signac, which uses MACS2 to call peaks[34]. We performed this peak calling by cell type, annotated according to the RNA-seq data, and these peak sets were then combined. Peaks found to be in genomic blacklist regions or on nonstandard chromosomes were removed.

To visualize the ATAC data we performed the standard Signac workflow with the following commands: RunTFIDF, FindTopFeatures, RunSVD, and RunUMAP. We then combined this representation of the data with the transcriptomic data visualization. This

representation of both RNA and ATAC data was created using the FindMultiModalNeighbors function in Signac, followed by running UMAP on the weighted nearest neighbor (WNN) graph.

Differential gene expression analysis

Before analysis, counts were normalized using the NormalizeData command in Seurat. Differential gene expression was performed on each cell type cluster using the R package presto[35]. Genes were considered significantly different if the adjusted p-value (Bonferroni correction) was less than 0.05. Comparisons were made between 5xFAD and WT mice at 2 months of age, 5xFAD and WT mice at 8 months of age, AD and cognitively non-impaired patient samples, and MCI and cognitively non-impaired patient samples.

CellChat and NeuronChat analysis

CellChat (v1.6.1)[10] and NeuronChat (v1.0.0)[16] were used to infer communication between cell types. First, CellChat and NeuronChat's standard workflows were run on each condition within the snRNA-seq data individually: 2-month-old 5xFAD mice, 2-month-old WT mice, 8-month-old 5xFAD mice, and 8-month-old WT mice. Separately, human samples were also put through this pipeline: AD, MCI, and control. Comparative analyses were then performed between 5xFAD and WT mice at 2 months of age and 8 months of age. AD and cognitively normal patient samples, and MCI and cognitively normal patient samples were also compared.

hdWGCNA, mouse and human comparison

To detect modules of genes that co-vary in a cell type-specific manner, high dimensional weighted gene co-expression network analysis (hdWGCNA) was performed using the hdWGCNA package[36]. hdWGCNA adapts the well-established WGCNA method[37] for snRNA-seq data. This analysis was performed on cell types where there were at least 1,000 nuclei in the mouse dataset: microglia, oligodendrocytes, OPCs, astrocytes, GABAergic neurons, and glutamatergic neurons.

Modules were detected in each cluster within the mouse snRNA-seq data, projected onto the human data, and then module preservation and module quality were evaluated. Modules found to be of high quality and highly preserved across species (zsummary.qual >10 and zsummary.pres > 10) were further investigated. Differential module eigengene analysis was performed to analyze differences in expression of each module between control and disease groups.

This analysis was performed, and results visualized using EnrichR (v3.178)[38,39] and hdWGCNA functions RunEnrichr, GetEnrichrTable, and EnrichrBarPlot sequentially. Enrichment analysis was done using the 2021 versions of the Gene Ontology databases[40,41].

DIRECT-NET analysis and visualization, mice

DIRECT-NET[42] was applied to the mouse 10X Genomics Multiome data to detect regulatory links specific to either 5xFAD or WT mice at 2 and 8 months of age and to construct gene regulatory networks implicated in AD. Promoters for each gene in the mouse genome were defined to be the 1,000 basepairs before the transcription start site of a gene. The DIRECT-NET tutorial was then followed, with analyses being applied to each genotype and time point individually first and then compared. First, cis-regulatory elements (CREs) were inferred following the DIRECT-NET tutorial for genes of interest based on previous analyses, including *Gria4*, *Grin2b*, and *Grm1*. Links were detected by genotype by subsetting the data and then running the Run_DIRECT_NET function on each genotype at each age. Signac's coverage plot function was used to visualize these links. Only high-confidence (HC) links were plotted, with links specific to 5xFAD mice highlighted in pink, links specific to control mice highlighted in blue, and links apparent in both genotypes highlighted in purple.

We then proceeded to construct gene regulatory networks based on genes upregulated in 5xFAD mice for both the 2- and 8-month time points. Following the DIRECT-NET tutorial, we

used presto[43] to determine genes that were upregulated in 5xFAD mice according to auROC analyses (auc > 0.5) of snRNA-seq data. We excluded genes corresponding to ribosomal proteins and those without known functions, then evaluated which of the remaining genes overlapped with regions of the genome that were more accessible in 5xFAD mice based on the accompanying snATAC-seq data. CREs relating to both upregulated and accessible genes were then used to generate links to known transcription factors using the generate_peak_TF_links command in DIRECT-NET. These links were then constructed into a network and visualized using the Python package networkx[44].

3.3 Results

Sample selection and quality control for mouse and human AD samples

One significant advantage of using rodent models of AD is that AD pathology can be studied across the lifespan, including at pre-symptomatic time points. We chose two time points for this study: 2 months, a pre-symptomatic time point when plaques begin to form in select regions in the brain, and 8 months, after significant plaque accumulation has occurred and cognitive deficits have emerged[22,23]. We used female mice only, as female 5xFAD mice show more rapid pathogenesis than males[45–47]. We similarly only evaluated tissue samples from female postmortem brain tissue, as women have an increased risk of developing AD compared to men[48]. We performed a multiomic assessment (single nucleus RNA-seq plus single nucleus ATAC-seq) of cells from the ENT of 5xFAD and age-matched controls from each age group (**Figure 3.1A**). After alignment and quality control assessments of data, we had 23,412 cells for the 2-month time point and 11,776 cells for the 8-month time point. All expected cell types were captured in each dataset, including excitatory neurons, inhibitory neurons, astrocytes, microglia, oligodendrocytes, and oligodendrocyte precursor cells (OPCs; **Figure 3.1B-C**). Cell types were similarly represented

across time points and genotypes (**Figure S3.1C-E**). To confirm that our integrated multiomic strategy worked as expected, we assessed the correspondence between the snRNA-seq and snATAC-seq datasets. Marker genes such as *Slc17a7*, *Mog*, and *C1qb* were more accessible and more highly expressed in glutamatergic neurons, oligodendrocytes, and microglia, respectively, than in other cell types (**Figure 3.1C-F**), consistent with known gene expression patterns. Accessibility of other marker genes for each cell type are shown in coverage plots in **Figure S3.2**.

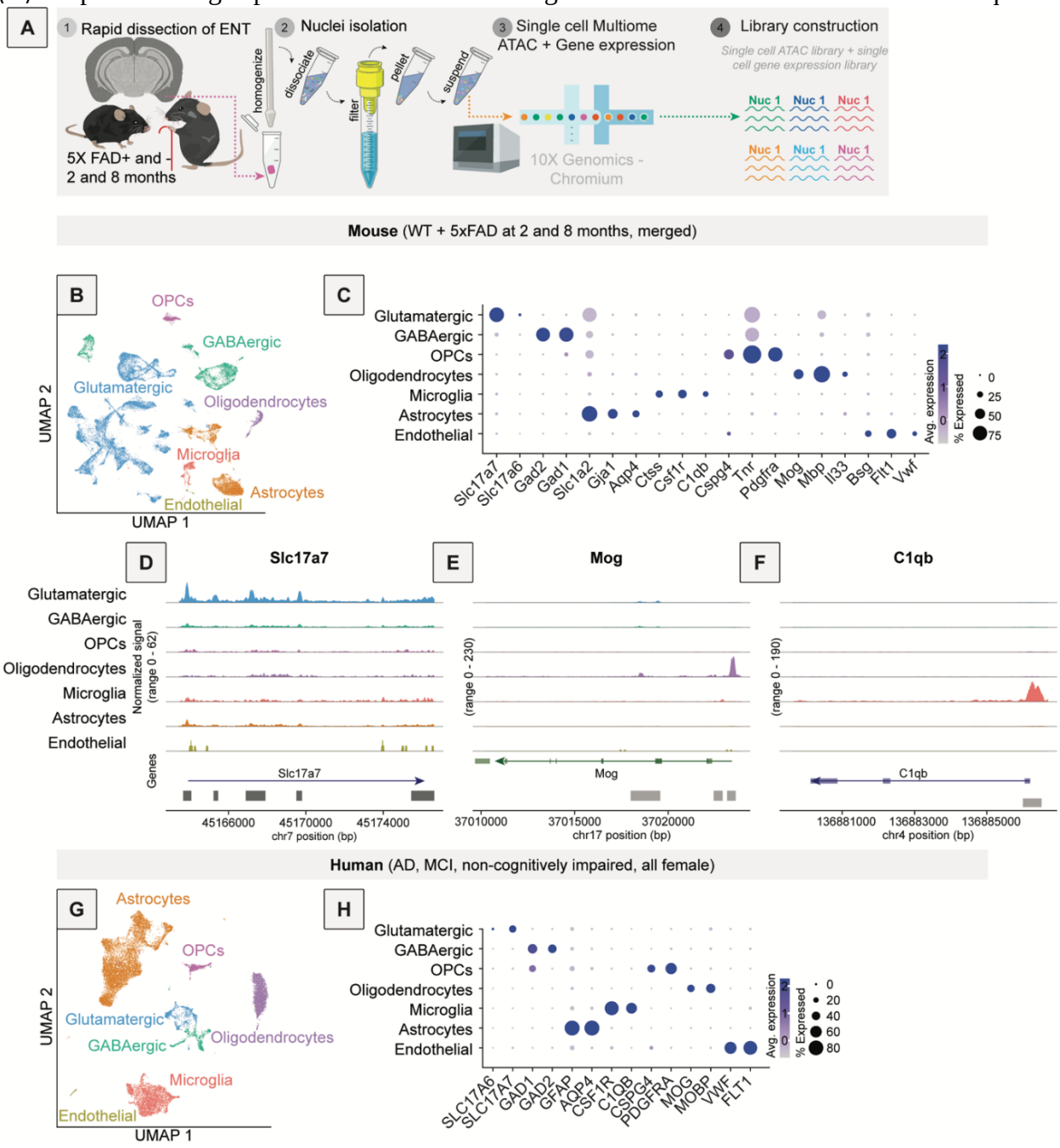
To explore the diversity of neuron types identified, we assessed the number of distinct subtypes of glutamatergic and GABAergic neurons within our mouse dataset. We performed unsupervised clustering using the smart local moving (SLM) algorithm and found 5 inhibitory (In) and 12 excitatory (Ex) neuronal subclusters with distinct transcriptional profiles (**Figure S3.3A**). We also transferred labels from a well-annotated dataset from the Allen Brain Institute that enabled definition of each cell type's layer in the brain and assessed gene expression differences (**Figure S3.3B-C**). Using this annotation method, we identified the following cell types: GABAergic subcluster 1 (In 1) was defined by expression of the vasal intestinal peptide (*Vip*), In 2 by synuclein gamma (*Sncg*), In 3 by lysosomal-associated membrane protein 5 (*Lamp5*), In 4 by parvalbumin (*Pvalb*), and In 5 by somatostatin (*Sst*). Excitatory neurons were classified by layer (depth in cortex, L2-6) and connectivity patterns (intratelencephalic, IT; corticothalamic, CT; pyramidal

Figure 3.1: Characterization of single cell sequencing data from mouse and human postmortem brain samples.
(A) Diagram of experimental workflow for snRNA-seq and snATAC-seq. The ENT of mice aged 2 months and 8 months, both 5xFAD and control, was dissected, nuclei were isolated, and sequencing performed using the Single Cell Multiome ATAC + Gene Expression kit from 10x Genomics.
(B) Single-nucleus data for mice aged 2 months and 8 months, both 5xFAD and control, represented in 2D UMAP (WNN) with general cell type annotations.
(C) Dotplot showing expression of known marker genes in each cell type in the mouse snRNA-seq portion of the multiome data.
(D-F) Coverage plots showing known marker gene accessibility as captured by the snATAC-seq portion of the multiome data. These data show that chromatin accessibility is reflective of gene

expression (i.e., the area near the transcription start site is more open in genes that are more highly expressed in particular cell types). Shown are coverage plots for *Slc17a7* (D), a marker for glutamatergic neurons; *Mog* (E), a marker for oligodendrocytes; and *C1qb* (F), a marker for microglia.

(G) UMAP of merged, re-processed and annotated snRNA-seq data from postmortem tissue samples obtained from AD, MCI, and non-cognitively impaired patient brains. This is a combination of previously published datasets from female donors[5,26], and one newly sequenced tissue sample.

(H) Dotplot showing expression of known marker genes in the annotated human snRNA-seq data.



tract, PT; near projecting, NP).

To assess how our mouse results compared with those from human postmortem brain tissue, we utilized data from previously published studies from ENT tissue in humans with and without AD. We also performed snRNA-seq of postmortem brain tissue from one female patient with mild cognitive impairment (MCI) to acquire data from an earlier stage of disease. After merging and processing the human data, we observed the same cell types as we found in our mouse data, though in different quantitative proportions (**Figure 3.1G**). Annotation of cell types in the human data was based on the same marker genes as used in the mouse data (**Figure 3.1H**, **3.1C**). As we identified fewer neurons in the human samples, we did not perform further annotation of neurons beyond inhibitory and excitatory.

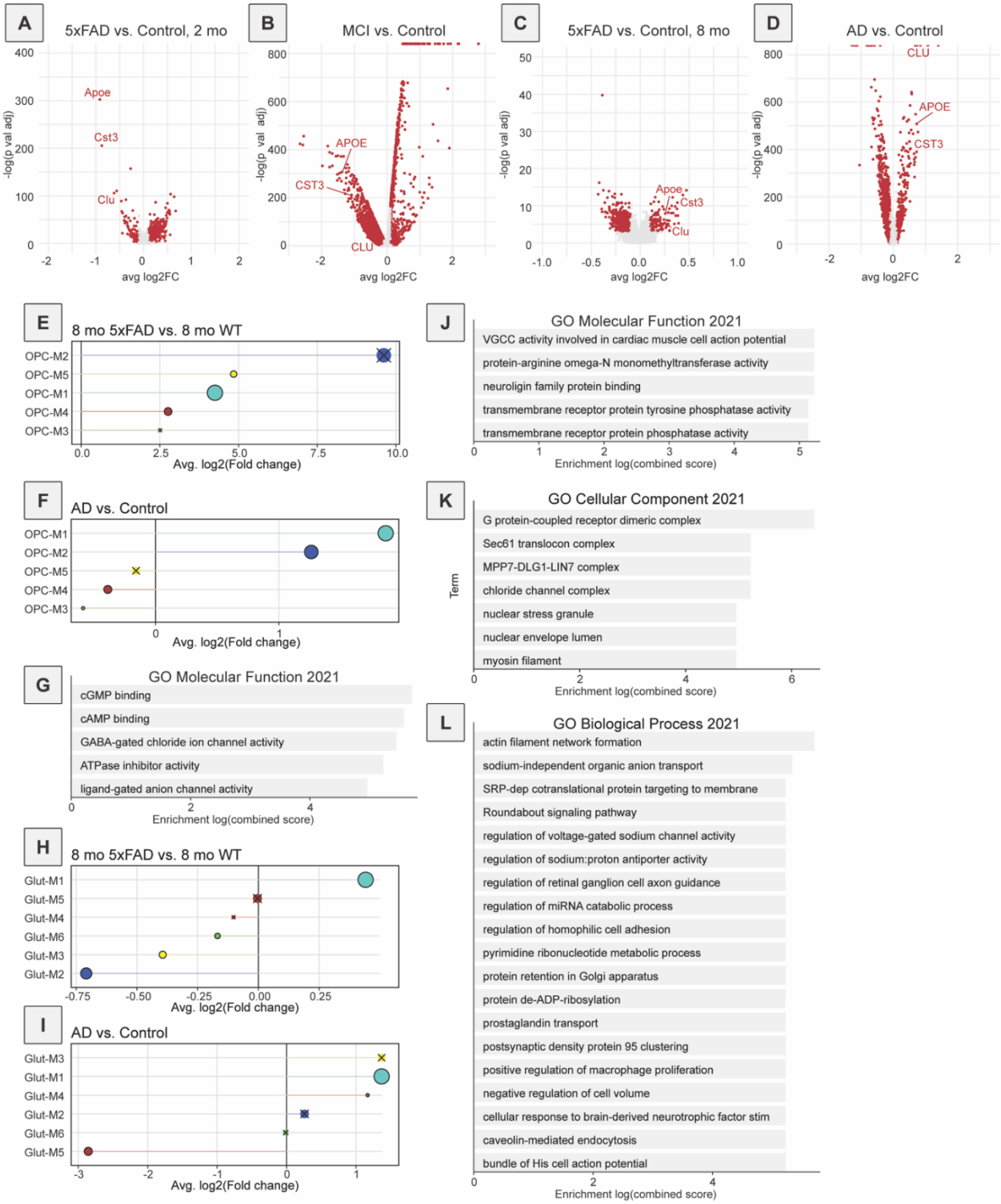
Differential gene expression and high dimensional weighted gene correlation network analysis (hdWGCNA) indicate a greater conservation of neuron-related gene expression changes relative to glial-related gene expression changes between mouse models and human AD

We next explored differences in gene expression between 5xFAD versus control mice, and human AD/MCI postmortem samples versus samples from non-cognitively impaired human postmortem brain tissue. Results of differential gene expression analysis for each cell type and each of the four comparisons (5xFAD versus control at 2 months, 5xFAD versus control at 8 months, MCI versus control, and AD versus control) including fold-change and adjusted p-value are reported in Supplementary Table 1, which will be available for download after publication.

As a benchmark, we first focused on several genes known to be associated with disease-associated astrocytes (DAAs) that emerge during the development of AD: Clusterin (*Clu*), Apolipoprotein E (*Apoe*), and Cystatin C (*Cst3*) [49,50]. DAAs have previously been identified in the

hippocampus and cortex of both human AD patients and 5xFAD mice[49,50]. Genes upregulated in DAAs are associated with amyloid clearance, aging, and the complement system[51]. Unexpectedly, at 2 months of age in 5xFAD mice, DAA genes were downregulated compared to controls (**Figure 3.2A**). Interestingly, we see a very similar downregulation of *APOE* and *CST3* in the human MCI brain (**Figure 3.2B**). However, in 5xFAD mice at 8 months of age the DAA genes *Apoe*, *Clu*, and *Cst3* were significantly upregulated, consistent with previous reports in mice and humans that these genes are enriched in DAAs or upregulated in AD in subclusters of astrocytes (**Figure**

Figure 3.2: Differential gene expression and conserved modules of gene expression changes from mouse and human data using hdWGCNA.
(A) Differential gene expression in astrocytes for 5xFAD versus control mice at 2 months of age. A positive log-fold change here indicates that a gene is upregulated in the 5xFAD genotype.
(B) Differential gene expression in astrocytes for human MCI vs non-cognitively impaired controls. A positive log-fold change here indicates that a gene is upregulated in MCI.
(C) Differential gene expression in astrocytes for 5xFAD versus control mice at 8 months of age. A positive log-fold change here indicates that a gene is upregulated in the 5xFAD genotype.
(D) Differential gene expression in astrocytes for human AD vs non-cognitively impaired controls. A positive log-fold change here indicates that a gene is upregulated in AD.
(E) DME analysis of OPC modules in 8-month-old 5xFAD mice. OPC M1 is significantly upregulated in 5xFAD mice at 8 months. If a module was not significantly different between diagnoses, an 'x' indicates an adjusted p-value of over 0.05.
(F) DME analysis of OPC modules in human AD compared to non-cognitively impaired individuals. The gene expression module OPC M1 is significantly upregulated in AD. If a module was not significantly different between diagnoses, an 'x' indicates an adjusted p-value of over 0.05.
(G) GO analysis of genes within OPC M1, from molecular function database. cGMP and cAMP binding are among the most prominent terms, as are those related to ion channels.
(H) DME analysis of Glutamatergic neuron modules in 8-month-old 5xFAD mice. Glut M1 is upregulated in 5xFAD mice at 8 months.
(I) DME analysis of glutamatergic neuron modules in human AD. Glut M1 is upregulated in human AD.
(J) GO analysis of genes within Glut M1, molecular function database. Two of the top terms relate to voltage-gated calcium channels, which control cellular excitability, and neuroligin-mediated cellular interactions.
(K) GO analysis of genes within Glut M1, cellular component database. G-protein coupled receptors, which can mediate a variety of cellular processes including excitability, and chloride channels are among the top terms.
(L) GO analysis of genes within Glut M1, biological process database. Highlights include genes relating to intrinsic neuronal excitability and cell adhesion.



3.2C)[49,52]. We observed a significant upregulation of each of these genes in the human AD brain (**Figure 3.2D**). These data together indicate that gene expression patterns occurring in the earliest stages of disease may not mirror those that occur at later stages, and in some cases may trend in

the opposite direction. This in turn may indicate that gene expression changes during AD pathogenesis may not progress in a linear fashion and could be dynamically regulated during disease.

While assessing changes in the expression of individual genes is useful to assess specific changes, a more holistic approach may better represent the overall patterns of gene expression changes during disease progression. Previously, weighted gene co-expression network analysis (WGCNA[37]) was applied to bulk RNA-seq data to detect such gene networks, which allowed for further analysis and comparison of these gene modules. This approach was recently modified for use with single-cell data (hdWGCNA[36]). Here we applied hdWGCNA to each cell population in the mouse dataset (oligodendrocytes, microglia, OPCs, astrocytes, GABAergic neurons, and glutamatergic neurons). We then projected these cell type-specific mouse gene modules onto the corresponding human gene expression data and quantified module preservation across species and module quality[53]. Only high-quality modules, meaning reproducible and robust, and highly preserved modules with a similar network architecture in both species were further examined. This quality and preservation analysis was done to focus our efforts on modules whose expression change was not due to chance, and was not unique to the mouse model to increase the likelihood that our results are relevant to human AD.

Interestingly, we found that in microglia and astrocytes, two cell types commonly studied in the context of AD in rodent models including 5xFAD[54–57], no modules were both highly preserved and of high quality (**Figure S3.4A-B**), indicating that patterns of gene expression changes were distinct in 5xFAD mice and human AD for these cell types. In oligodendrocytes just one module, Oligo-M3, was both highly preserved and of high quality (**Figure S3.4C**). To evaluate whether expression of genes in this module was significantly different between genotypes, we used

differential module eigengene (DME) analysis implemented in hdWGCNA and found that the module expression was significantly upregulated in 2-month-old 5xFAD mice and in postmortem brain samples from both human MCI and AD compared to samples from non-cognitively impaired patients (**Figure S3.4D, F-G**). In 8-month-old 5xFAD mice a similar effect was seen but it was not statistically significant (**Figure S3.4E**). Similarly, in GABAergic neurons one module, GABA-M3, was both highly preserved and of high quality (**Figure S3.5A**). This module was not significantly differentially expressed between samples from 5xFAD and control mice at either age, nor postmortem brain samples from AD or MCI patients and controls (**Figure S3.5B-E**), suggesting that the differences in expression were relatively mild. In OPCs, one module, OPC-M1, was both highly preserved and of high quality (**Figure S3.5F**). The downregulation of genes in this module was not significant at 2 months (**Figure S3.5G**), however, this module was downregulated in MCI (**Figure S3.5H**) but upregulated in human AD and in 5xFAD mice (**Figure 3.2E-F**). We further investigated what genes were expressed in OPC-M1 and to what pathways they may contribute. OPCs primarily differentiate into oligodendrocytes, but there is also evidence that they can become neurons or astrocytes under certain conditions[58–61]. OPCs are involved in biological processes including myelination[62] as well as synaptic remodeling and pruning[63–65]. Based on gene ontology (GO) analysis, genes in OPC-M1 were most highly associated with cGMP and cAMP binding (**Figure 3.2G**). cGMP and cAMP are both well-known small molecule second messengers for a variety of signal transduction pathways and have both been shown to modulate synaptic plasticity[66–68]. Additionally, it has previously been suggested that cAMP and cGMP alter metabolism of amyloid-β or stimulate its production[69–71], which is especially relevant given amyloid-β's abundance in the 5xFAD mouse brain.

Overall, the best module quality and conservation was observed in glutamatergic neurons, where we observed 2 modules, GluT-M1 and GluT-M3, that were both highly preserved and of high quality (**Figure S3.5I**). DME analysis indicated that only GluT-M1 was significantly differentially expressed – upregulated – in both 5xFAD mice at both time points and postmortem human AD brains relative to their respective controls (**Figure 3.2H-I**, **Figure S3.5J-K**). Differences were not significant in MCI. GO analysis of this gene module highlighted biological processes including voltage-gated calcium channel activity (**Figure 3.2J**), G protein-coupled receptors (GPCRs; **Figure 3.2K**), and actin filament network formation (**Figure 3.2L**), all processes that are important in synaptic regulation[72–74]. Interestingly, though the GABA-M3 module noted previously was not significantly differentially expressed, GO analysis of GABA-M3 also yielded terms associated with calcium transport and glutamate receptors (**Figure S3.5L**), indicating a common set of neuronal changes relating to cellular excitability. Together, these results suggest that gene expression patterns related to neuron-specific signaling processes, including synaptic modulation by OPCs, intrinsic cellular excitability, and glutamatergic signaling are better preserved between 5xFAD mice and human AD brains than modules specific to glial cells, and that these neuronal processes may play a central role in disease pathophysiology.

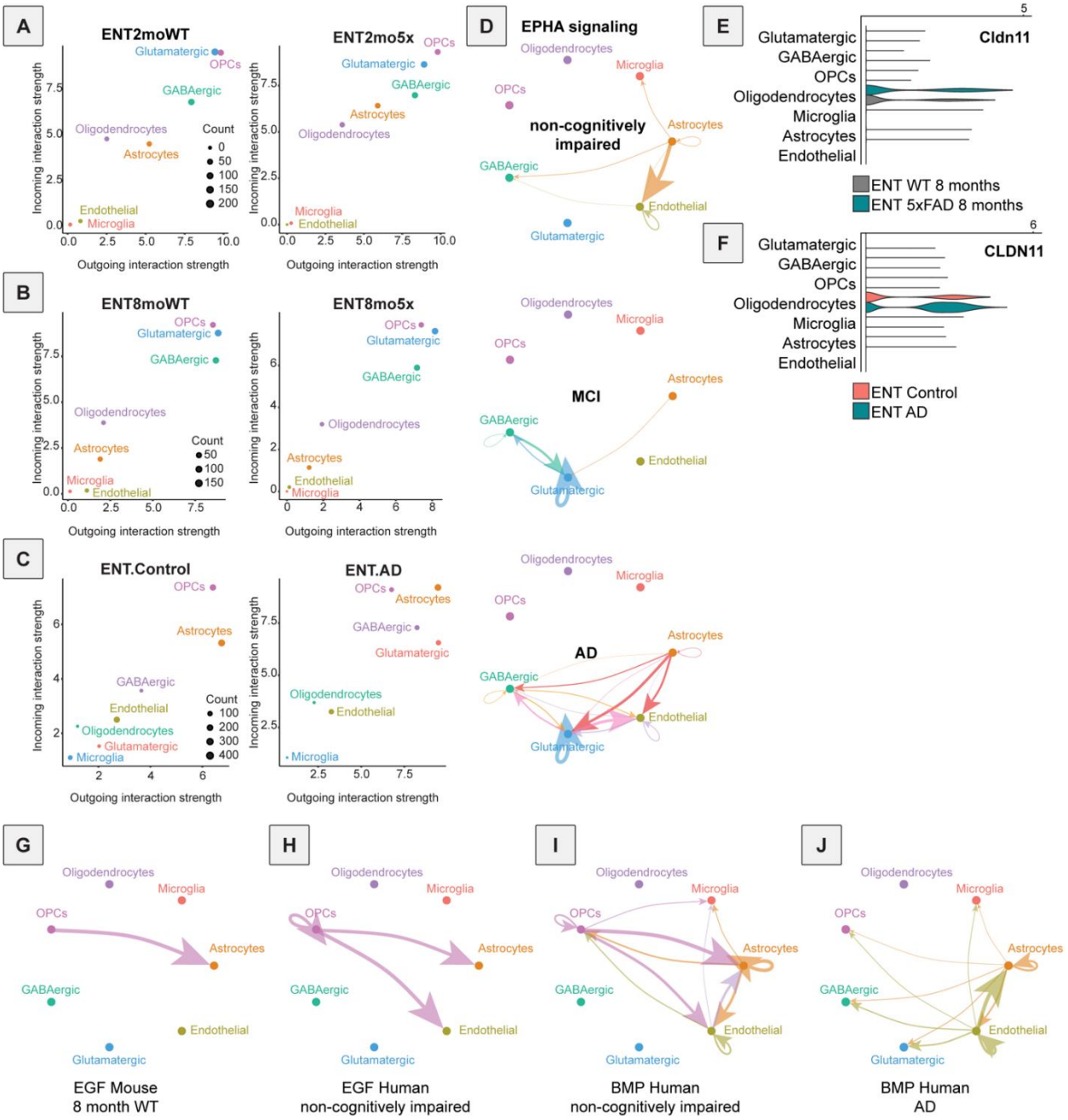
Changes in inter-cellular communication networks during AD development

While single cell sequencing data enable the study of gene expression and gene module regulation within defined cell types, one key limitation is that these results only focus on cell-intrinsic changes. This provides only a partial picture of disease, as the brain consists of numerous cell types that operate in a dynamic fashion during both health and disease, and appropriate inter-cellular communication is necessary for proper brain function. For example, as glial cells play critical roles in synaptic function and remodeling which occur in neurons, focusing only on the

transcriptome of glial cells in isolation, and not considering how glial changes impact communication with neurons, may mask their true importance in disease. To gain a deeper understanding of the evolution of inter-cellular interactions within the ENT, we utilized the computational package CellChat[75]. This method infers changes in cell-cell communication networks based on known cellular communication pathways. CellChat determines potentially important signaling pathways by identifying in which cell populations known ligands and corresponding receptor genes are expressed, including multimeric proteins, as well as other proteins that may be involved. Based on expression of these genes, we can infer modes of inter-cellular communication, which can be visualized as a directed network where nodes are cell types and edges are connectivity probabilities. These networks can inform us of likely communication pathways between identified cell types, and how the strength of these predicted interactions change across conditions and/or disease.

We first assessed the overall inferred level of incoming and outgoing signaling for each cell type, based on snRNA-seq data. In both 2- and 8-month-old control mice, the majority of inter-cellular signaling – both outgoing and incoming – was predicted to either arise from or be received

Figure 3.3: Application of CellChat reveals common and species-specific changes in cellular communication in mouse and postmortem human brain tissue.
(A) Predicted outgoing and incoming cellular signaling with CellChat for each cell type in the mouse ENT at 2 months of age for control mice (left) and 5xFAD mice (right).
(B) Predicted outgoing and incoming cellular signaling with CellChat for each cell type in the mouse ENT at 8 months of age for control mice (left) and 5xFAD mice (right).
(C) Predicted outgoing and incoming cellular signaling with CellChat for each cell type in the human ENT for patients without cognitive impairment (left) and patients with AD (right).
(D) CellChat signaling inferred for EPHA in patients (from top to bottom) without cognitive impairment, with MCI, and with AD.
(E) Violin plot of *Cldn11* expression at 8 months of age in mice.
(F) Violin plot of *Cldn11* expression in AD patients, and those without cognitive impairment.
(G) CellChat signaling inferred for EGF in mouse, control, 8 months of age.
(H) CellChat signaling inferred for EGF from brain tissue donors without cognitive impairment.
(I) CellChat signaling inferred for BMP in samples from patients without cognitive impairment.
(J) CellChat signaling inferred for BMP from brain tissue donors with AD.



by OPCs or neurons, with microglia and endothelial cells predicted to contribute the least to inter-cellular signaling (**Figure 3.3A-B**). Notably, only relatively small differences were detected in 5xFAD mice relative to controls when comparing data from 2- and 8-month-old animals, primarily in astrocytes and oligodendrocytes (**Figure 3.3A-B**), indicating only subtle changes in inferred inter-cellular communication occur over time in 5xFAD mice. In contrast, while we observed similar elevations in predicted signaling in astrocytes and oligodendrocytes in human MCI and

AD tissue samples, changes in predicted signaling in glutamate and GABA cells was much more pronounced, with more outgoing and incoming signaling predicted in each cell type in MCI and AD than controls (**Figure 3.3C, S3.6A**). These data together indicate that 5xFAD mice may exhibit similar patterns of signaling changes as in human AD, though to a lesser extent (**Figure 3.3A-C, S3.6A**).

We next identified molecular signaling pathways that showed similar or different changes between 5xFAD mice and human AD samples (**Figure S3.6B**). Pathways that exhibited the most similar disease-related changes across species were Ephrin-A (EPHA), Claudin (CLDN), Epidermal Growth Factor (EGF), and Bone morphogenetic protein (BMP). Predicted signaling through EPHA pathways was highly conserved between mice and humans across time points, and there was a higher level of predicted signaling in the EPHA pathway in 5xFAD mice at both timepoints as well as human MCI and AD compared to controls (**Figure 3.3D**, **S3.6B-D**). Glutamatergic neurons showed the largest increase in predicted EPHA signaling in the disease state relative to controls, an effect that was more pronounced in human MCI and AD than 5xFAD mice (**Figure 3.3D**, **S3.6B-D**). While the EPHA signaling pathway includes the genes *Epha1-8* and *Efna1-5*, this elevation in inferred signaling was mostly driven by expression changes in *Epha4* and *Epha7*, which were significantly upregulated in glutamatergic neurons in AD compared to controls (*EPHA4*, logFC = 0.695, adjusted p-value = 0.000557; *EPHA7*, logFC = 0.386, adjusted p-value = 0.00468, Supplementary Table 1). Interestingly, a loss of EphA4 was previously shown to improve aspects of social memory[76] and synaptic plasticity in the hippocampus[77], suggesting that an upregulation in EphA4, as observed here in AD-related disease, may impair properties of neuronal signaling. Interestingly, other studies have also linked EphA4 to amyloid-β, though the nature of this link is not clear and may be bidirectional. For example, while some studies suggested

that Epha4 regulates amyloid-β[78], others indicated that amyloid-β regulates *Epha4*[77]. While our results do not provide an answer to this discrepancy, they support the observation that as amyloid-β accumulates, *Epha4* expression increases.

Notably, while glutamatergic neurons showed the largest predicted differences in EPHA signaling, the *Epha3*, *Epha4*, *Epha5*, and *Epha7* genes were also present in the high-quality and highly preserved OPC-M1 discussed previously (Supplementary Table 2, available for download after publication). In addition, the two GO terms previously mentioned as highly associated with OPC-M1 – cAMP and cGMP signaling – are engaged in the EPHA signaling pathway[79], and are upregulated, consistent with an elevation in signaling through EPHA. Therefore, an upregulation in EPHA signaling may impact multiple cell types and contribute to neuronal dysfunction during the development of AD-related disease.

In the CLDN signaling pathway, only one gene was implicated, *Cldn11*, which was significantly upregulated in both 8-month-old 5xFAD mice and postmortem human AD tissue samples (**Figure 3.3E-F**). We found that the upregulation of *Cldn11* in oligodendrocytes is opposite of that found in a previous study performed in human postmortem AD samples, where a downregulation was observed[80]. Notably, that study was performed using tissue from the precuneus region in the brain rather than the ENT, and therefore, one possibility is that changes may be brain region-dependent. Supporting this interpretation, another study using the amyloidogenic APP-NL-G-F[81] mice found that a gene module containing *Cldn11* was either upregulated or downregulated depending on the brain region under study[82].

Predicted EGF signaling was also different in both 5xFAD mice and postmortem human AD samples relative to controls. Samples from control mice and non-cognitively impaired postmortem brain samples exhibited predicted EGF signaling from OPCs to astrocytes (**Figure**

3.3G-H). However, no EGF signaling was predicted in 5xFAD mice at 8 months of age, or tissue samples from MCI or AD patients. Notably, there was still predicted EGF signaling in 2-month-old 5xFAD mice, consistent with these animals representing an early disease stage (**Figure S3.6B**, **E**). Regarding the potential function of EGF, a previous study showed that EGF treatment prevented cognitive impairment in female mice, which is consistent with the hypothesis that a loss of EGF signaling contributes to AD[83]. Therefore, our data suggest that signaling through the EGF pathway is reduced in AD, consistent with the protective effect of restoring EGF signaling in diseased brains.

Finally, BMP signaling was also predicted to exhibit similar changes in 5xFAD mice and human AD brains (**Figure 3.3I-J**, **Figure S3.6F-G**). While no significant BMP signaling was predicted in control or 5xFAD mice at 2 months of age, BMP signaling was predicted to occur at 8 months of age, and slightly more so in controls than 5xFAD mice (**Figure S3.6F-G**). The difference was mostly manifest as a reduction in astrocyte-astrocyte signaling during disease progression (**Figure S3.6F-G**). Similarly, in human postmortem brain samples, BMP signaling was predicted to be stronger in controls compared to samples of patients with either MCI or AD (**Figure 3.3I-J**, **S3.6B**). In this case, changes in astrocyte-astrocyte signaling were still predicted, though the most substantial change was predicted to occur in OPCs, in which predicted signaling was reduced in AD (**Figure 3.3I-J**). Genes implicated in these changes include *Bmpr2*, *Bmpr1b*, *Bmpr1a*, and *Bmp7* (**Figure S3.6H**). While BMP signaling genes *Bmp4* and *Bmp6* have previously been linked to AD and were shown to be upregulated in other mouse models of AD, expression of these two genes was not significantly different in controls vs. disease samples in our mouse data[84] (Supplementary Table 1). In the human data, *Bmp7* expression was downregulated in AD compared to control samples in astrocytes (logFC = -0.217, adjusted p-value = 8.94E-78,

Supplementary Table 1). Notably, *Bmp7*, in contrast to *Bmp4* and *Bmp6*, has a potentially protective effect in AD[85]. Thus, the reduction in *Bmp7* signaling parallels the observed reduction of EGF signaling in AD, as both are thought to be forms of neurotropic signaling and protective against disease. These results thus point to a potential loss of trophic signaling as a common feature of AD pathogenesis in mice and humans.

In addition to pathways where similar inferences were made between mice and humans, we also detected some pathways that had contrasting predicted changes between species. For example, PTPRM had higher predicted signaling in brains from AD patients but lower in 5xFAD mice, PSAP was higher in AD patients but lower in 5xFAD mice, and PDGF was lower in AD patients and higher in 8-month-old 5xFAD mice relative to controls (**Figure S3.6B**). PTPRM has been suggested to play a role synapse formation[86], and though it has not been widely studied in AD, a previous Genome-Wide Association Study identified two SNPs within the PTPRM gene associated with dementia[87]. As these SNPs would not be present in the 5xFAD model, this may be an explanation for the discrepancy between species. In contrast, PSAP, a lysosomal trafficking protein[88], has been proposed as a potential biomarker for preclinical AD[89], and a drug targeting PSAP reduced neuronal loss and decreased inflammation in a mouse disease model[90]. The PDGF signaling pathway has also long been linked to AD[91], with decreases in PDGF signaling being associated with disruption of the blood brain barrier and eventual neurodegeneration[92]. It is possible that these differences across species reflect differences in disease staging, or they may reflect larger differences in disease between the rodent 5xFAD model and human AD.

NeuronChat predicts a conserved increase in glutamatergic transmission in the 5xFAD rodent model and human AD

Our CellChat analysis provided a set of inter-cellular signaling processes that change during disease, data which implicated changes in signaling between both neurons and glia. This provides a global picture of how inter-cellular signaling may be disrupted in the ENT during the development of AD. One limitation of the CellChat package and database is that it only considers classic ligand-receptor interactions between cells where both ligand and receptor are expressed proteins. However, most neuron-neuron and some neuron-glia communication occurs via small molecule neurotransmitters such as glutamate and GABA. Therefore, this analysis may miss a major component of disease-related changes that develop during AD. To explore potential changes in neurotransmitter-based communication, we applied the newly developed computational package NeuronChat, which utilizes a similar analysis method to CellChat but is adapted for neurotransmitter-based communication[93]. A focus on disease in neurons is warranted, as of the five hdWGCNA gene modules found to be both highly preserved and of high quality, three were found in neuronal populations, and one other in OPCs which have neuron-like features. This raises the possibility that exploring changes in inter-neuronal signaling in 5xFAD mice may have the most relevance for the human AD brain.

We used NeuronChat to compute all the predicted changes in neurotransmitter-based communication in cells from mice and human postmortem brain samples (**Figure 3.4A-E**). Changes were observed in 5xFAD mice: at 2 months of age, there was an increase in signaling to oligodendrocytes and endothelial cells (**Figure S3.6I-J**). In 8-month-old 5xFAD mice there was an increase in neurotransmitter-based communications, occurring approximately equally in both GABAergic and glutamatergic cells (**Figure 3.4A-B**, **E**). In human postmortem AD brain samples, most changes in signaling were predicted to occur from glutamatergic neurons, with a general increase in predicted communication from glutamatergic neurons to both GABA and glutamatergic

neurons in the AD brain (**Figure 3.4C-D**). Several predicted increased interactions were observed in both 5xFAD mice at 8 months of age and human AD, including between glutamate and AMPA receptors (Glu-Gria1, Glu-Gria2, Glu-Gria4), NMDA receptors (Glu-Grin2b and Glu-Grin1), kainate receptors (Glu-Grik2), metabotropic glutamate receptors (Glu-Grm1, Glu-Grm5, Glu-Grm8), and the neuronal adhesion molecules neurexin and neuroligin (Nrxn3_Nlgn1; **Figure 3.4E**). Notably, no changes in GABAergic signaling pathways were predicted. In addition, genes mediating interactions – *Gria2*, *Grin1*, *Grin2b*, *Grm1*, and *Grm5*w – were also present in the preserved hdWGCNA module Glut-M1 (Supplementary Table 2), indicating that expression of genes for these receptors varied across cells in a coordinated fashion across species. When considering genes implicated in glutamate production rather than receptors, we found that *Gls* was significantly upregulated in glutamatergic neurons in 8-month-old 5xFAD mice and tissue samples from human AD brains (logFC = 0.0594, adjusted p-value = 4.79E-05, and logFC = 0.427, adjusted p-value = 4.49E-14, Supplementary Table 1). *Gls* was downregulated in 2-month-old 5xFAD mice (logFC = -0.177, adjusted p-value = 7.45E-71, Supplementary Table 1), and not significantly altered in MCI. *Slc17a7*, or vesicular glutamate transporter 1 (vGluT1), was also upregulated in glutamatergic neurons from both 8-month-old 5xFAD mice and human AD brain samples (logFC = 0.0999, adjusted p-value = 2.95E-10, and logFC = 0.271, adjusted p-value = 0.0464, Supplementary Table 1). Similarly to *Gls*, it was downregulated in 2-month-old 5xFAD mice (logFC = -0.0777, adjusted p-value = 5.46E-13, Supplementary Table 1), and not significantly differentially expressed in MCI.

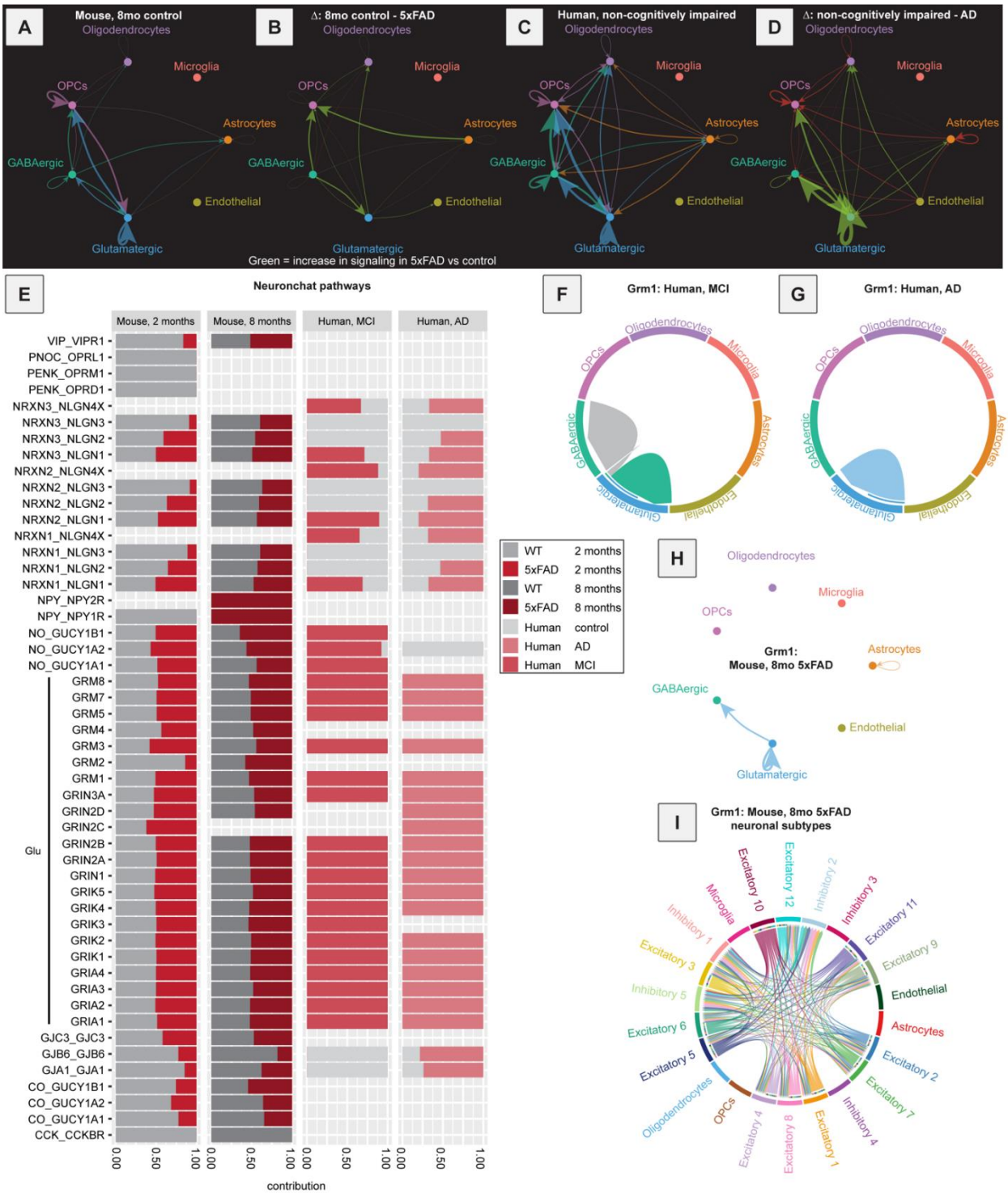
Related to increased glutamatergic signaling in AD, there is a long-held hypothesis in AD that excitotoxicity mediated by excessive glutamatergic signaling is a key contributor to disease pathogenesis[94–96]. Our combined observations that 1) most predicted changes in neurotransmitter-

based interactions occurred via glutamatergic signaling, 2) glutamatergic neurons had two gene expression modules that were both high quality and highly preserved, more than any other cell type, 3) the EPHA signaling pathway is predicted to strongly impact glutamatergic neurons all support this hypothesis. As many of the genes relevant to glutamatergic signaling (*Gria2*, *Grin1*, *Grin2b*, *Grm1*, and *Grm5*) were part of Glut-M1 as identified by hdWGCNA, we were curious what other genes may also be in this module. As expected from the GO analysis shown in **Figure 3.2J-L**, many ion channel genes were found that contribute to intrinsic neuronal excitability, including sodium ion channel genes *Scn1a*, *Scn2a*, *Scn3a*, *Scn3b*, *Scn8a*, and calcium ion channel genes *Cacna1e*, *Cacna1b*, *Cacna1c*, *Cacna2d1*, and *Cacng3* (Supplementary Table 2). Of these genes, *Cacna1b* and *Cacna1c* were significantly upregulated in human AD, and *Cacna1c* was significantly upregulated in 5xFAD at 8 months of age (Supplementary Table 1), positioning this gene to play a conserved and potentially key role in AD. Upregulation of these genes would likely lead to an elevation in intrinsic neuronal excitability, though it is not clear if these changes may be drivers of pathogenesis, or compensatory responses to disease development. *Cacna1c* has previously been linked to bipolar disorder and depression[97], and more recently to AD using exploratory bioinformatic analyses[98]. In mice, an increase in *Cacna1c* expression has been negatively correlated with object recognition memory[99]. Interestingly, its expression in reactive astrocytes has also been associated with plaque formation, providing a potential mechanistic role for *Cacna1c* in AD pathogenesis[100].

We next looked more closely at one particular pathway, the glutamate to metabotropic glutamate receptor 1 (Glu-Grm1, a G_q-coupled GPCR that activates phospholipase C), because 1) Glu-Grm1 was within the highly conserved and high-quality Glut-M1 module, 2) it was identified as differentially regulated in NeuronChat analyses, and 3) it was upregulated in glutamatergic

neurons in 5xFAD mice at 8 months of age (logFC = 0.161, adjusted p-value = 4.49E-14, Supplementary Table 1). Unsurprisingly, in both 8-month-old 5xFAD mice and human postmortem brain samples from patients with MCI or AD, glutamatergic neurons were the predicted sender of this Glu-Grm1 interaction in all cases (**Figure 3.4F-H**). To define which neuron types may be involved in the predicted Glu-Grm1 interaction, we examined communication in specific annotated neuronal cell clusters of 8-month-old 5xFAD mice (**Figure 3.4I**). Every excitatory cluster (subtypes 1-12) exhibited at least some predicted outbound communication via Glu-Grm1, as expected, and inbound glutamate was also widespread, except for in subtype 10 in both 5xFAD and control mice. Interestingly, for GABAergic interneurons, inbound Glu-Grm1 interactions were only predicted onto some subtypes; for example, GABAergic subclusters inhibitory 3, characterized by *Lamp5* expression (**Figure S3.3B-C**), and inhibitory 4, characterized by *Pvalb* expression (**Figure S3.3B-C**), did not have any significant predicted Grm1-Glu communication (**Figure 3.4I**), which reflects low expression of *Grm1* in these subclusters. These

Figure 3.4: NeuronChat analysis of mouse and human data detail an elevation of glutamatergic signaling during AD progression.
(A) Signaling inferred for all NeuronChat pathways detected in 8-month-old control mice, aggregated together by signaling strength.
(B) Differences in signaling inferred for all NeuronChat pathways aggregated together by signaling strength in 5xFAD versus control mice at 8 months of age. An increase in signaling in 5xFAD vs control is indicated in green, and a decrease in signaling in red.
(C) Signaling inferred for all NeuronChat pathways detected in non-cognitively impaired patient data, aggregated together by signaling strength.
(D) Differences in signaling inferred for all NeuronChat pathways aggregated together by signaling strength in AD versus non-cognitively impaired patients. An increase in signaling in AD vs control is indicated in green, and a decrease in signaling in red.
(E) Comparison of predicted signaling for all NeuronChat pathways in 5xFAD versus control mice at 2 and 8 months of age (left), and in patients with MCI or AD versus patients without cognitive impairment (right)
(F) Signaling inferred via *Grm1* in AD patients between different cell types.
(G) Signaling inferred for *Grm1* in MCI patients.
(H) Signaling inferred for *Grm1* in 8-month-old 5xFAD mice.
(I) Same data as (H), broken down by individual neuronal subtypes.



results suggest that elevations in glutamatergic signaling may differentially impact select subtypes of glutamatergic and GABAergic inhibitory neurons in the ENT.

In addition to glutamatergic signaling, we also found differences in predicted interactions via neuronal cell adhesion molecules. For example, we observed in NeuronChat analyses a predicted increase in communication via Neurexin-3 and Neuroligin-1 in human MCI and AD, though this was not seen in mice (Nrxn3-Nlgn1, **Figure 3.4E**). Similar to what we found in data from human postmortem brain samples, expression of *Nlgn1* was significantly upregulated in glutamatergic neurons in 5xFAD mice at 8 months of age (log2FC = 0.163, adjusted p-value = 7.66E-16, Supplementary Table 1), and *Nrxn3* was upregulated in astrocytes, though this was not statistically significant (log2FC = 0.148, adjusted p-value = 0.357, Supplementary Table 1). Neuroligins are post-synaptic cell-adhesion molecules on excitatory synapses[101], and expression of *Nlgn1* has previously been shown to induce increases in glutamatergic synapse density and activity[102]. We also observed an increase in the expression of neuroligin-related binding proteins in the GO term from GluT M1 (**Figure 3.2J**). Thus, the observed upregulation of *Nlgn1* is consistent with an elevation in excitatory synaptic communication, which would be expected to contribute to excitotoxicity. *Nrxn3* is a pre-synaptic protein which can bind to neuroligins, among other proteins, and has been found to be downregulated in postmortem brain samples from AD patients, though these samples were from the hippocampus[103]. Indeed, the function of *Nrxn3* is likely brain region-dependent[104]. Taken together, our results indicate that increased glutamatergic signaling in AD occurs via changes in multiple cell types including glutamatergic neurons but also GABAergic neurons, OPCs, and astrocytes, consistent with a system-wide disease state.

Disease-associated epigenetic regulation reflects an elevation in glutamatergic communication

The observed systems-wide remodeling of communication between both neurons and glia in the brain are likely dependent on epigenetic modifications that occur during disease. In our

rodent sequencing experiments, we performed a multiomic assessment (combined snRNA-seq and snATAC-seq) which enables us to link chromatin accessibility to gene expression across all cell types.

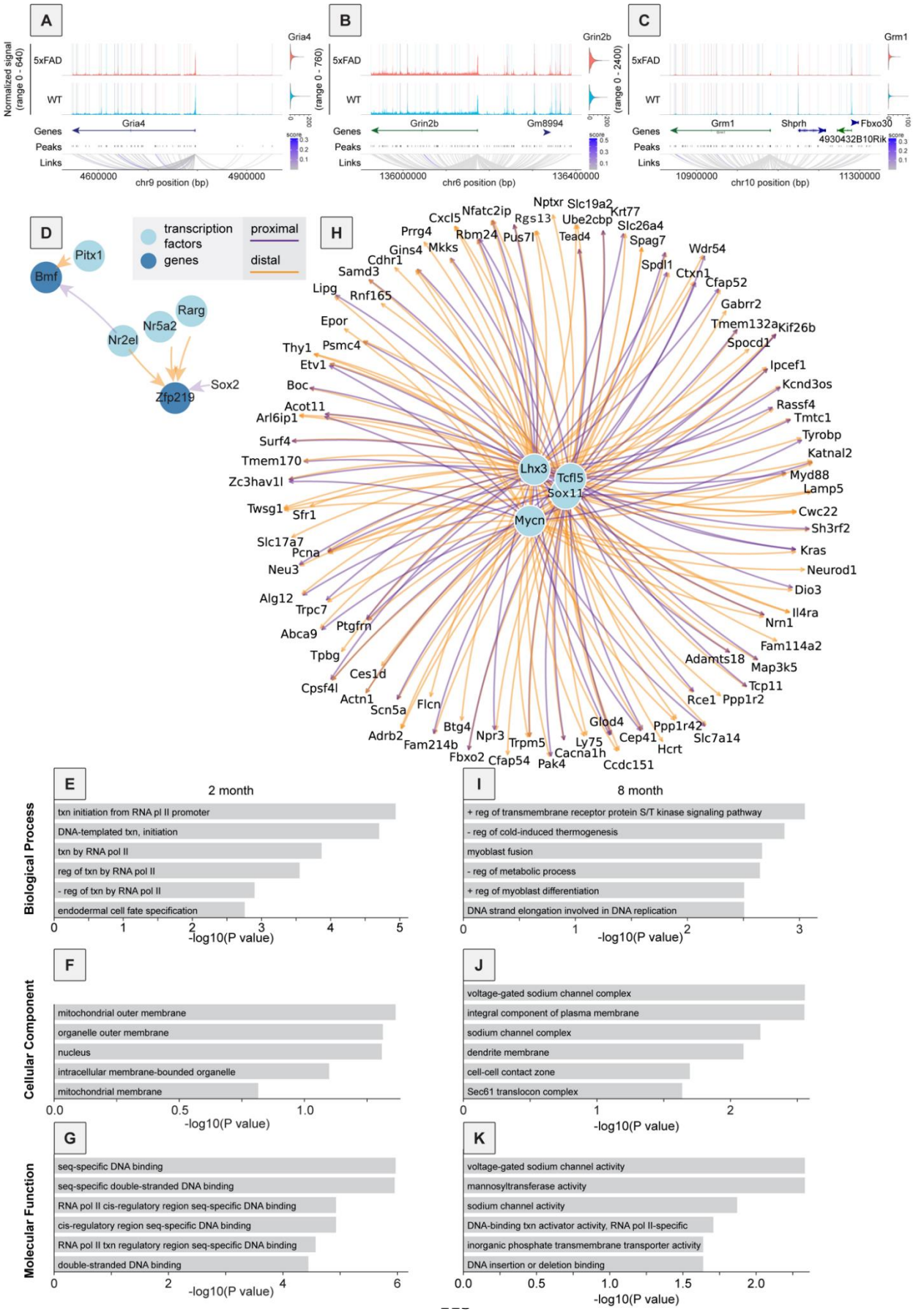
To assess the relationship between chromatin accessibility and gene expression, we utilized DIRECT-NET, which identifies both cis-regulatory and trans-regulatory elements within cellular DNA. Cis-regulatory elements are detected using a machine learning method that links peaks within the snATAC-seq data to gene expression values reflected in the snRNA-seq data, whereas trans-regulatory elements are detected by matching sequences within peaks to known transcription factors. DIRECT-NET was run on data from 5xFAD and control mice from 2-month-old and 8-month-old groups, run separately. This was done so that if links between peaks to specific genes were only found in one genotype or age, we could isolate in which conditions these links were identified. We focused our initial analyses on genes identified in our NeuronChat analyses. Results showed some links specific to 5xFAD, some specific to control mice, and some present in both genotypes (**Figure 3.5A-C**). We found that genes encoding glutamate receptors such as *Gria4* and *Grin2b* had many high confidence (HC) links to peaks in the ATAC-seq data, but none in coding regions of the genome outside of the gene of interest itself (**Figure 3.5A-B**). In contrast, *Gria4* had 2 HC links outside of the *Gria4* gene locus, which may indicate that these regions are important in the expression of *Gria4*, though they were not specific to either WT or 5xFAD mice, so we did not pursue these regions further. *Grin2b* had more links to loci external to coding regions with some being specific to 5xFAD or control mice.

Of particular interest was *Grm1*, as in mice at 8 months of age, *Grm1* was upregulated in 5xFAD mice relative to controls (logFC = 0.161, adjusted p-value = 4.49E-14, Supplementary Table 1) and had multiple linked peaks specific to each genotype. In addition to non-coding loci,

the expression of this gene was linked to expression of multiple other genes. For example, we observed a link specific in 5xFAD mice from *Grm1* to the transcription start site of the gene *Shprh* (**Figure 3.5C**). To our knowledge, *Shprh* has not previously been linked to AD, though it is known to be highly expressed in the human brain and a potential tumor suppressor gene[105]. There was also a link to the transcription start site of the E3 ubiquitin ligase Fbxo30, which was specific to control mice and not observed in 5xFAD mice. Interestingly, Fbxo30 has previously been implicated in the positive regulation of BMP signaling[106], which we noted in our CellChat analyses in **Figure 3.3** and **S3.6**. While Glu-Grm1 interactions identified by NeuronChat were limited to glutamatergic neurons and certain GABAergic neurons, changes in BMP signaling were predicted to occur between astrocytes and other cell types (e.g., **Figure 3.3I-J**, **S3.6F-G**). This finding linking *Grm1* expression to genes that control BMP signaling suggests that this may be another route, albeit indirect, by which neurons and glial cells may interact.

We next proceeded to expand our multiomic assessment to any gene that was both

Figure 3.5: Multi-omics approach reveals links between glutamatergic signaling via *Grm1* and trophic signaling via BMPs.
(A-C) Coverage plots of mouse ATAC-seq data with peaks highlighted according to what genotype linkages were detected in using DIRECT-NET. Blue indicates linkage to the gene of interest in control mice only. Red indicates linkage to the gene of interest in 5xFAD mice only. Purple indicates linkage to the gene of interest in both genotypes.
(A) Coverage plot of *Gria4* and the surrounding genome area with links to the *Gria4* TSS. A violin plot of *Gria4* RNA expression is shown to the right.
(B) Coverage plot of *Grin2b* and surrounding genome area with links to the *Grin2b* TSS.
(C) Coverage plot of *Grm1* and surrounding genome area with links to the Grm1 TSS.
(D) GRN resulting from DIRECT-NET analysis of genes upregulated in 5xFAD mice at 2 months of age.
(E-G) GO analysis of genes in the GRN obtained in Figure 3.5D for biological processes (E), cellular component (F), and molecular function (G).
(H) GRN resulting from DIRECT-NET analysis of genes upregulated in 5xFAD mice at 8 months of age.
(I-K) GO analysis of genes in the GRN obtained in Figure 3.5H for biological processes (I), cellular component (J), and molecular function (K).


A
B
C
Normalized signal (range 0 - 640)
5xFAD
WT
Gria4
Genes
Peaks
Links
4600000
chr9 position (bp)
4900000
Grin2b
(range 0 - 760)
Gm8994
136000000
chr6 position (bp)
136400000
(range 0 - 2400)
Grm1
Shprh
Fbxo30
4930432B10Rik
10900000
chr10 position (bp)
11300000
D
transcription factors
genes
proximal
distal
Pitx1
Bmf
Rarg
Nr5a2
Nr2el
Sox2
Zfp219
H
Nptxr
Slc19a2
Nfatc2ip
Cxcl5
Rgs13
Ube2cbp
Krt77
Slc26a4
Prrg4
Rbm24
Pus7l
Tead4
Spag7
Wdr54
Gins4
Mkks
Cdhr1
Spdl1
Ctxn1
Samd3
Cfap52
Lipg
Rnf165
Gabrr2
Epor
Tmem132a
Kif26b
Spocd1
Thy1
Psmc4
Ipcef1
Etv1
Kcnd3os
Boc
Rassf4
Acot11
Tmtc1
Arl6ip1
Tyrobp
Surf4
Katnal2
Tmem170
Lhx3
Tcfl5
Myd88
Zc3hav1l
Sox11
Lamp5
Twsg1
Sfr1
Cwc22
Mycn
Sh3rf2
Slc17a7
Pcna
Kras
Neu3
Neurod1
Dio3
Alg12
Trpc7
Il4ra
Nrn1
Abca9
Ptgfrn
Fam114a2
Tpbg
Adamts18
Map3k5
Ces1d
Tcp11
Cpsf4l
Actn1
Rce1
Ppp1r2
Scn5a
Flcn
Glod4
Ppp1r42
Adrb2
Btg4
Npr3
Trpm5
Ly75
Cep41
Slc7a14
Fam214b
Cacna1h
Hcrt
Fbxo2
Cfap54
Pak4
Ccdc151
E
2 month
Biological Process
txn initiation from RNA pl II promoter
DNA-templated txn, initiation
txn by RNA pol II
reg of txn by RNA pol II
- reg of txn by RNA pol II
endodermal cell fate specification
-log10(P value)
I
8 month
+ reg of transmembrane receptor protein S/T kinase signaling pathway
- reg of cold-induced thermogenesis
myoblast fusion
- reg of metabolic process
+ reg of myoblast differentiation
DNA strand elongation involved in DNA replication
-log10(P value)
F
Cellular Component
mitochondrial outer membrane
organelle outer membrane
nucleus
intracellular membrane-bounded organelle
mitochondrial membrane
-log10(P value)
J
voltage-gated sodium channel complex
integral component of plasma membrane
sodium channel complex
dendrite membrane
cell-cell contact zone
Sec61 translocon complex
-log10(P value)
G
Molecular Function
seq-specific DNA binding
seq-specific double-stranded DNA binding
RNA pol II cis-regulatory region seq-specific DNA binding
cis-regulatory region seq-specific DNA binding
RNA pol II txn regulatory region seq-specific DNA binding
double-stranded DNA binding
-log10(P value)
K
voltage-gated sodium channel activity
mannosyltransferase activity
sodium channel activity
DNA-binding txn activator activity, RNA pol II-specific
inorganic phosphate transmembrane transporter activity
DNA insertion or deletion binding
-log10(P value)

upregulated in 5xFAD mice according to our snRNA-seq data, and any loci differentially accessible according to our snATAC-seq data. In 2-month-old mice, DIRECT-NET analyses resulted in a small gene regulatory network (GRN) which included transcription factors Zfp219 and Bmf (**Figure 3.5D**). This network's small size was largely because at 2 months of age, very few loci were significantly differentially accessible between control and 5xFAD mice, which is expected given the early time point in disease progression. To characterize genes within the network we performed a GO analysis and found, unsurprisingly given the role of Zfp219 as a transcription factor, that genes in the network were highly associated with transcription and DNA binding (**Figure 3.5E-G**).

Results from this same analysis method but at 8 months of age yielded a much larger GRN (**Figure 3.5H**), consistent with the later disease stage, resulting in larger and more substantial changes in chromatin accessibility and gene expression regulation. Overall, this GRN contained 4 transcription factors and 89 genes. GO analysis on this set of genes showed not only association with regulation of transcription and DNA binding, as expected, but also association with ion channel activity (**Figure 3.5I-K**). Specifically, voltage-gated sodium channel activity and sodium channel activity were among the top 3 terms relating to molecular function (**Figure 3.5K**), another indication that regulation of ion channels may be a principal feature of disease, both at the gene expression and chromatin regulation levels. Notably, one of these transcription factors within the network, Sox11, has previously been implicated in controlling intrinsic neuronal excitability, with its expression being activity-regulated[107], raising the possibility that this gene may play an important role in regulating the elevation of intrinsic neuronal excitability that occurs in AD.

3.4 Discussion

Here, we show that while inferred signaling changes during AD are widespread between different cell types in the ENT in both 5xFAD mice and human AD, changes in glutamatergic signaling are the most prominent and shared feature of disease in mice and postmortem human samples. This conclusion is supported by several lines of evidence. First, two of the five total gene expression modules identified using hdWGCNA that were both highly conserved and high-quality were found in glutamatergic neurons, while a third was found in GABAergic neurons, and a fourth in OPCs, which all have processes related to neuronal excitability and/or neuronal synapses (**Figure 3.2**, **S3.5**). In addition, we found using CellChat that the major signaling changes occurred in, glutamatergic cells, GABAergic cells, and OPCs both early and late in disease (**Figure 3.3**). Changes in predicted EPHA signaling in AD mostly occurred in glutamatergic neurons (**Figure 3.3D**), but multiple genes associated with EPHA signaling were also observed in the OPC gene expression module M1, which also contained the GO terms cAMP- and cGMP-binding, which is critical for EPHA signaling (**Figure 3.2G**, Supplementary Table 2). Lastly, our analysis focusing on neuron-specific modes of communication using NeuronChat indicated an elevation in glutamatergic signaling as the main feature of AD pathogenesis in both mice and humans, including the up-regulation of *Slc17a7*, *Gls*, and *Grm1* (**Figure 3.4**, Supplementary Table 1). Regulation of *Grm1* was linked to *Fbxo30*, which in turn regulates BMP signaling. As both *Grm1* and BMP signaling processes are reduced in AD, this may be a mechanism by which changes in glutamatergic signaling coordinate with other cell populations to trigger wider modifications in signaling between cell types.

CellChat and NeuronChat implicate tissue-wide changes in cellular communication during AD

The advent of single cell RNA sequencing technologies has enabled an unprecedented definition of the molecular changes that occur in a variety of disease processes, including in AD. Indeed, there has been a substantial increase in the number of recent publications detailing the molecular changes that occur during AD in a variety of rodent models and human postmortem tissue samples, at multiple time points and in multiple brain regions[5,6,8,26,108–112]. However, one core limitation of these technologies is that they only explore cell-autonomous changes that occur during disease, and do not consider interactions between cells. Thus, traditional sequencing analyses provide only a partial picture of the disease, making it difficult to explore tissue- and systems-level questions. A previous study utilized CellChat to explore the changes in cellular communication that occur between non-neuronal cells[113]; however, how this relates to neuron-neuron and neuron-glia interactions is only just starting to be characterized in humans[114], and cross-species comparisons are yet to explored. Here we define the gene expression changes in putative receptor-ligand interactions that are conserved between mouse and humans, and focus on neuron-specific changes in communication, identifying changes in glutamatergic signaling across multiple cell types and signaling pathways. For example, the major cell types that are predicted to engage in inter-cellular signaling are glutamatergic neurons, GABAergic neurons, and OPCs (**Figure 3.3A-C**). Given that these cells together contain 4 of the 5 gene expression modules that are well-preserved and high-quality between mouse and human, these findings further support the need to understand how changes in communication between these cells contribute to disease progression. In the future, our findings will be further enhanced by the application of spatial transcriptomic technologies, which will allow us to add a spatial dimension to the changes in cellular communication that we observed here. We have recently developed several computational packages designed to integrate snRNA-seq and spatial transcriptomic datasets, including

COMMOT[115], which should further allow us to link changes in gene expression to spatial patterns of cellular communication, as well as pathology *in situ* including amyloid plaques and tau tangles.

One of the common changes between the 5xFAD model and postmortem human AD brains was decreased BMP signaling. This was seen early in the 5xFAD mouse model at 2 months of age and continued through 8 months. In data from human postmortem samples, expression was reduced in both MCI and AD. *Bmp7*, which has not been the focus of many previous studies in the context of AD[116], stood out as being both highly expressed and significantly differentially expressed across conditions, particularly in astrocytes and OPCs (**Figure S3.6H**, Supplementary table 1). BMP signaling has many previously established functions, including neurogenesis and astrogliogenesis[116], and *Bmp6* has even been shown to localize to amyloid-b plaques in the hippocampus (*Bmp7* was not tested)[117]. *Bmp7* specifically has been suggested to be neuroprotective, but of particular interest to us, it has been shown to reduce the negative effect of excess glutamate[118]. The fact that it is downregulated early in 5xFAD mice and in human AD suggests a reduced ability to respond to increased glutamate load, even before disease is severe.

Further cementing the importance of excitotoxicity were our findings regarding the EGF signaling pathway, which was predicted to decrease in 5xFAD samples at 8 months of age compared to controls, and in MCI and AD samples compared to controls. While EGF signaling has a multitude of functions, it, like *Bmp7*, has also been suggested to be neuroprotective, specifically against excess glutamate[119]. Most prominently, findings from NeuronChat indicated that increased glutamatergic signaling was a shared feature between 5xFAD and postmortem AD human tissue samples. While the glutamate receptor *Grm1* was significantly upregulated in 5xFAD mice, this was not the case in AD (Supplementary table 1), despite the NeuronChat results predicting increased signaling via *Grm1*. These observations suggest that perhaps in AD,

elevations in the production of glutamate, rather than its receptors, may be driving increased signaling. In addition, we found that the major changes in epigenetic regulation that occur in the 5xFAD model are related to cellular excitability and regulation of synaptic inputs. Indeed, genes that were identified as being part of the disrupted gene regulatory network in 5xFAD mice at 8 months of age largely related to sodium channels, which are a key contributor to excitotoxicity.

Neuronal excitability and excitotoxicity as the key conserved features of AD in mice and men

We were surprised to find that no gene expression modules in microglia or astrocytes were both well-preserved and high quality between 5xFAD mice and human AD, but that at least one gene expression module from each of the other cell types analyzed with hdWGCNA were significantly different (**Figure S3.4A-C, S3.5A, F, I**). This means that the coordinated gene expression changes that occur in microglia and astrocytes during disease development in 5xFAD mice are distinct from those that occur in human AD. If this is the case, it means that processes occurring in 5xFAD mice may not translate well to those occurring in humans. However, before making broad conclusions, several limitations to our study should be considered. First, our study only includes samples from the mouse and human ENT; thus, it is not clear if our findings may extend to other brain regions. Second, we only examined two time points in the mouse, at 2 and 8 months of age, and only females. It is possible that if we examined more time points, different time points, or both males and females, that we may observe better conservation in some of these groups. Third, levels of gene expression detected using snRNA-seq may not directly translate to protein levels, and the relationship between RNA and protein may be cell type-dependent.

Despite these limitations, at face value, our results suggest that the set of features of human AD that are best modeled in the 5xFAD model may be changes in neuronal signaling processes and intrinsic excitability, rather than glial cell processes. It is perhaps not surprising that only

certain features of human disease are well-modeled in rodents. For example, AD is a complex disease that includes, among other features, progressive aggregation of amyloid and hyperphosphorylated tau. However, most mouse models of AD pathogenesis only model one core feature of AD, for example amyloidogenesis in the 5xFAD mice. Therefore, molecular and cellular responses related to amyloid plaque aggregation and clearance in human AD may be well-recapitulated in the 5xFAD mouse, whereas molecular and cellular responses related to tau may not be. This, however, also raises the more general problem of how useful rodent models of AD are in studying human AD. No rodent model will perfectly capture all the dimensions of human AD, but perhaps we only need them to recapitulate the essential features of AD for the models to be sufficiently useful for therapeutic development. Indeed, the excitotoxicity hypothesis of AD has existed for many decades, though it has somewhat fallen out of favor, replaced in large part by the amyloid hypothesis, as well as involvement of other processes, such as the immune system[120,121]. It is possible that all of these are correct and interrelated; perhaps amyloid deposition triggers engagement of glia and inflammatory processes, which then effect changes in neuronal signaling that then drive excitotoxicity. If this is true, it is even more critically important to understand how changes in inter-cellular communication between glia and neurons may contribute to disease development and severity. Furthermore, it is important to leverage the strengths of the mouse lines that are used, and acknowledge their limitations, so that we can best focus on pathways and features of disease that are conserved between the mouse and human to maximize the translatability of our findings.

3.5 Supplementary figures

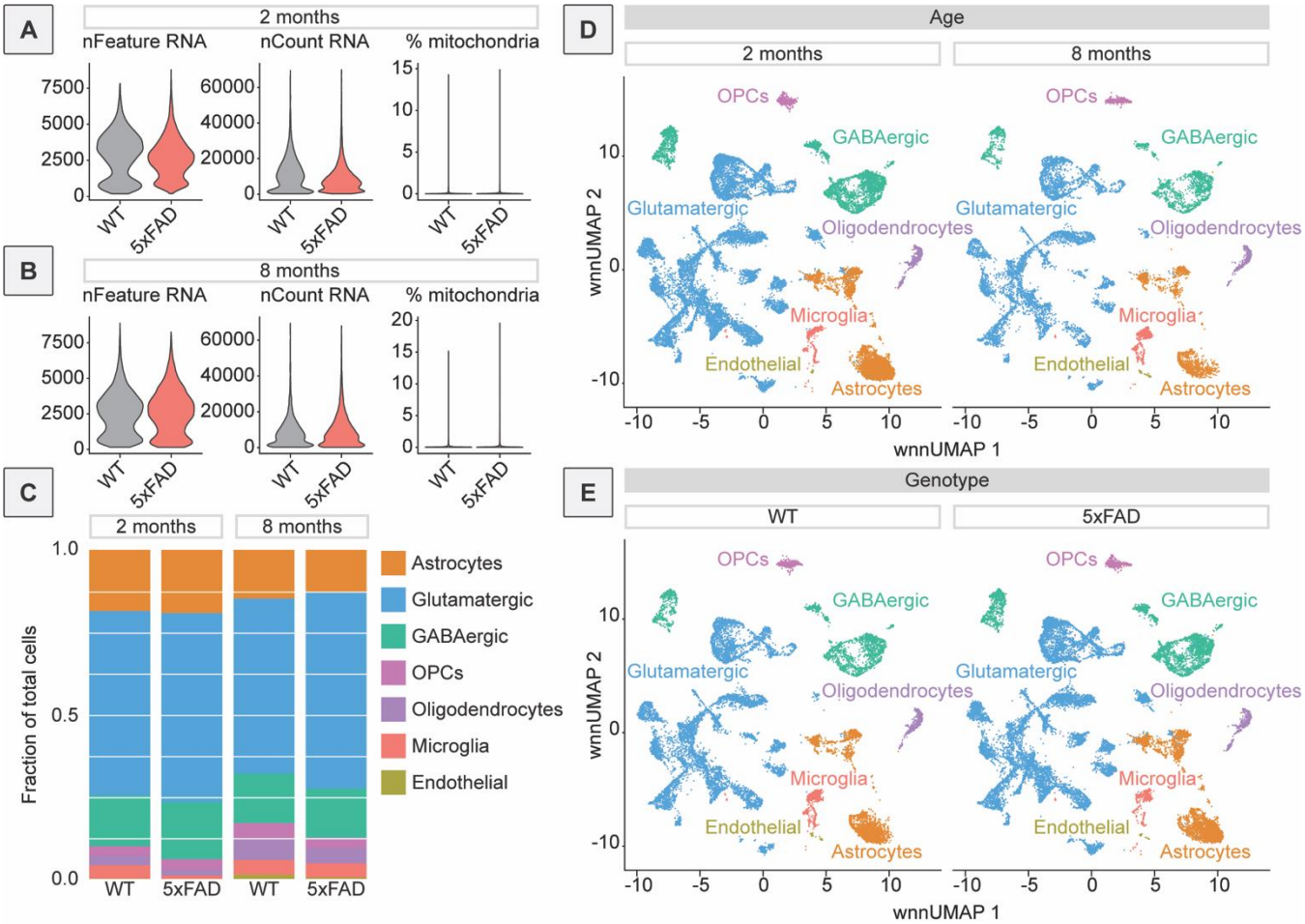
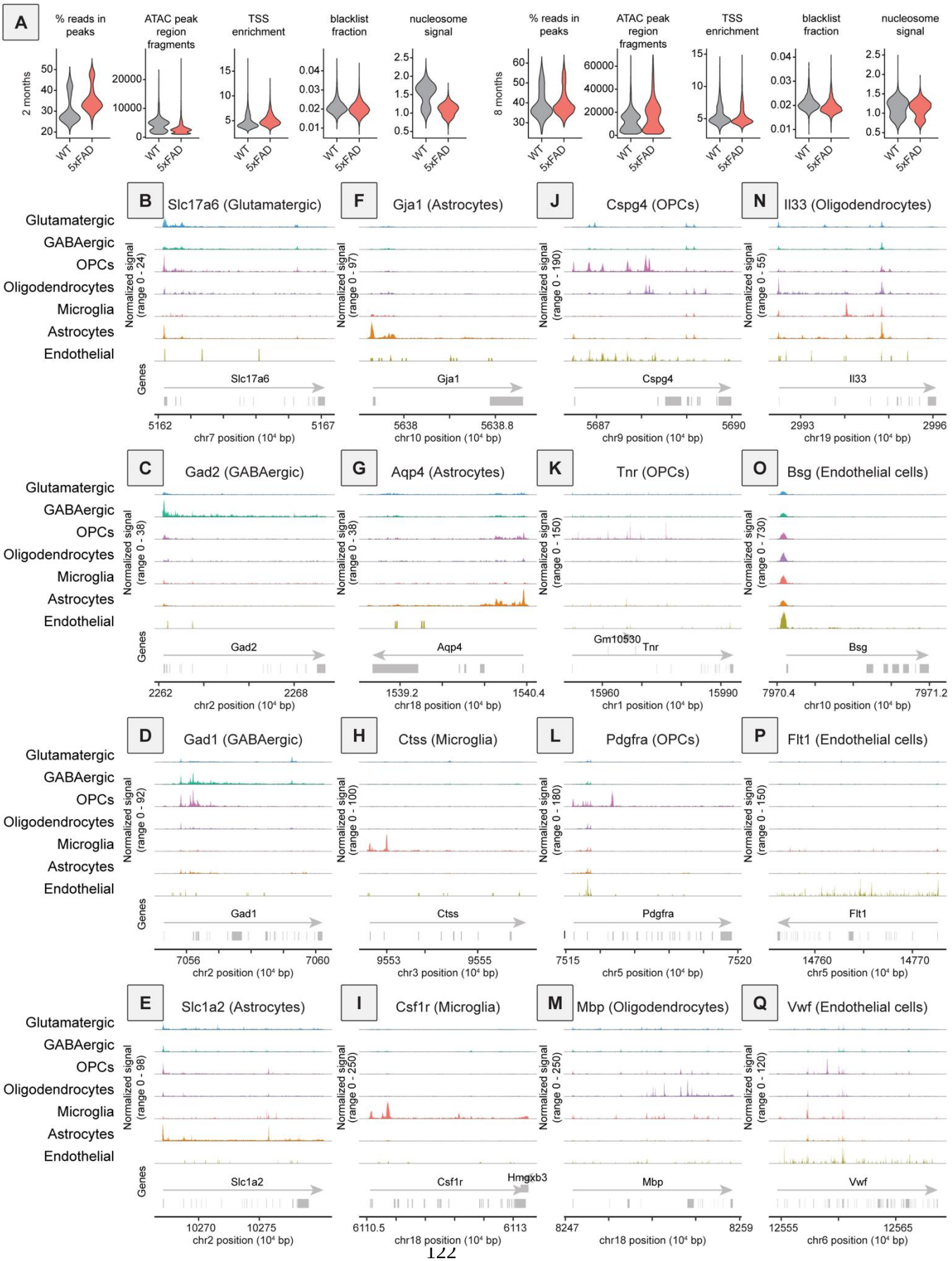


Figure S3.1: Additional quality control metrics for sequencing data.
(A) Violin plots of QC metrics for snRNA-seq data after filtering out low-quality cells. Data from 5xFAD and control mice at 2 months of age are plotted. Metrics included from left to right are nFeature RNA (number of genes detected), nCount RNA (number of molecules detected), and percent mitochondria (percent mitochondrial genes).
(B) Violin plots of QC metrics for snRNA-seq data after filtering out low-quality cells. Data from 5xFAD and control mice at 8 months of age are plotted.
(C) Bar plots showing cell type proportions for 5xFAD and control mice in the sequencing data.
(D) UMAP visualization of mouse sequencing data, split by age.
(E) UMAP visualization of mouse sequencing data, split by genotype.

Figure S3.2: Quality control metrics for snATAC-seq data, and example gene accessibility plots for various cell types in mouse ENT.
(A) Violin plots of QC metrics for snATAC-seq data after filtering out undesirable cells. ATAC QC metrics shown for 5xFAD and control mice at 2 (left) and 8 (right) months of age. From left to right are percent reads in peaks, ATAC peak region fragments, transcription start site enrichment, fraction of reads in blacklist regions, and nucleosome signal.
(B) Coverage plot for *Slc17a6*, a marker gene for glutamatergic neurons.
(C) *Gad2*, a marker gene for GABAergic neurons.
(D) *Gad1*, a marker gene for GABAergic neurons.
(E) *Slc1a2*, a marker gene for astrocytes.
(F) *Gja1*, a marker gene for astrocytes.
(G) *Aqp4*, a marker gene for astrocytes.
(H) *Ctss*, a marker gene for microglia.
(I) *Csf1r*, a marker gene for microglia.
(J) *Cspg4*, a marker gene for OPCs.
(K) *Tnr*, a marker gene for OPCs.
(L) *Pdgfra*, a marker gene for OPCs.
(M) *Mbp*, a marker gene for oligodendrocytes.
(N) *Il33*, a marker gene for oligodendrocytes.
(O) *Bsg*, a marker gene for endothelial cells.
(P) *Flt1*, a marker gene for endothelial cells.
(Q) *Vwf*, a marker gene for endothelial cells.


A
% reads in peaks
ATAC peak region fragments
TSS enrichment
blacklist fraction
nucleosome signal
2 months
8 months
WT
5xFAD
B Slc17a6 (Glutamatergic)
F Gja1 (Astrocytes)
J Cspg4 (OPCs)
N Il33 (Oligodendrocytes)
C Gad2 (GABAergic)
G Aqp4 (Astrocytes)
K Tnr (OPCs)
O Bsg (Endothelial cells)
D Gad1 (GABAergic)
H Ctss (Microglia)
L Pdgfra (OPCs)
P Flt1 (Endothelial cells)
E Slc1a2 (Astrocytes)
I Csf1r (Microglia)
M Mbp (Oligodendrocytes)
Q Vwf (Endothelial cells)
Glutamatergic
GABAergic
OPCs
Oligodendrocytes
Microglia
Astrocytes
Endothelial
Normalized signal
Genes
Slc17a6
Gja1
Cspg4
Il33
Gad2
Aqp4
Gm10530
Tnr
Bsg
Gad1
Ctss
Pdgfra
Flt1
Slc1a2
Csf1r
Hmgxb3
Mbp
Vwf
chr7 position (10⁴ bp)
chr10 position (10⁴ bp)
chr9 position (10⁴ bp)
chr19 position (10⁴ bp)
chr2 position (10⁴ bp)
chr18 position (10⁴ bp)
chr1 position (10⁴ bp)
chr10 position (10⁴ bp)
chr2 position (10⁴ bp)
chr3 position (10⁴ bp)
chr5 position (10⁴ bp)
chr5 position (10⁴ bp)
chr2 position (10⁴ bp)
chr18 position (10⁴ bp)
chr18 position (10⁴ bp)
chr6 position (10⁴ bp)

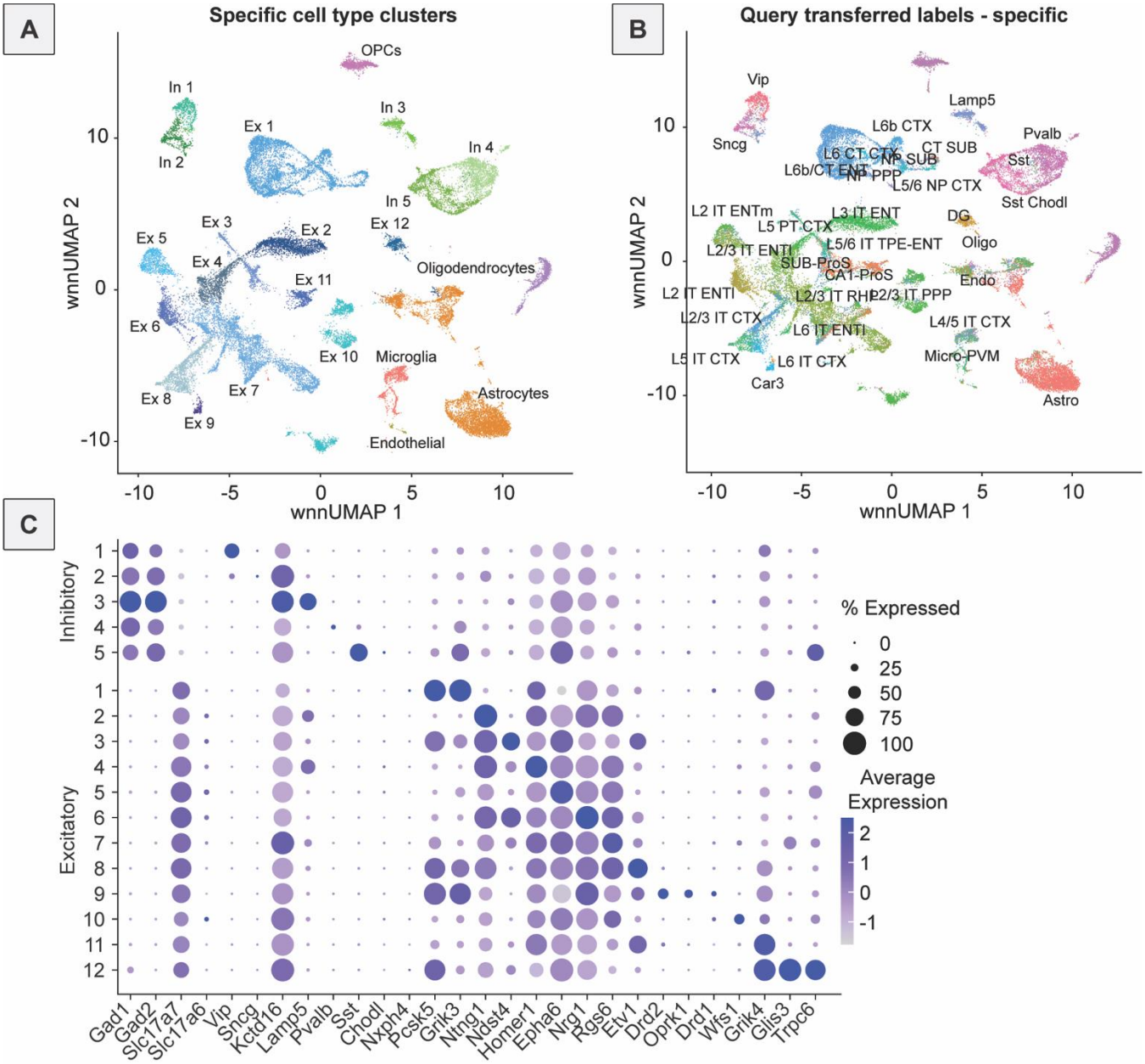


Figure S3.3: Sub-clustering and further annotation of neurons in 5xFAD and control mouse data.

(A) UMAP visualization of mouse sequencing data with subclusters for different types of GABAergic, or inhibitory (In), neurons and glutamatergic, or excitatory (Ex), neurons.

(B) UMAP visualization of mouse sequencing data annotated based on label transfer from data from the Allen Brain Atlas, including specific neuron types.

(C) Dot plot showing gene expression differences and expression of known marker genes in the subclustered neuronal data as labeled in Figure S3.3A.

Figure S3.4: hdWGCNA results for microglia, astrocytes, and oligodendrocytes.
(A) zSummary metrics indicating module quality and preservation across species for modules found within the microglia cluster using hdWGCNA.
(B) zSummary metrics indicating module quality and preservation across species for modules found within the astrocyte cluster using hdWGCNA.
(C) zSummary metrics indicating module quality and preservation across species for modules found within the oligodendrocyte cluster using hdWGCNA.
(D) DME analysis of oligodendrocyte modules in 2-month-old 5xFAD mice. All modules were upregulated in 5xFAD mice at 2 months of age.
(E) DME analysis of oligodendrocyte modules in 8-month-old 5xFAD mice. No modules were significantly different between genotypes at 8 months of age.
(F) DME analysis of oligodendrocyte modules in patients with MCI versus those without cognitive impairment.
(G) DME analysis of oligodendrocyte modules in patients with AD versus those without cognitive impairment.

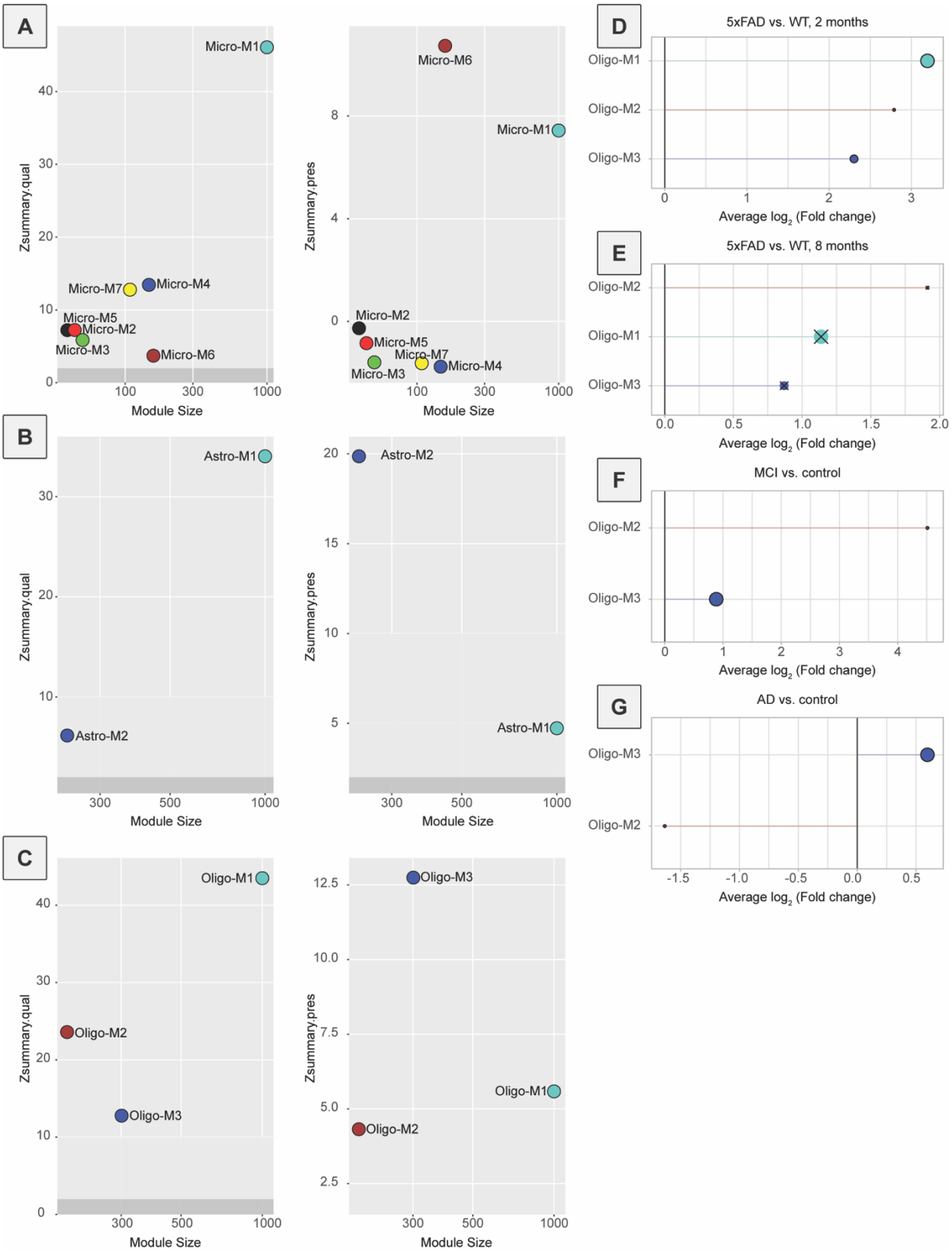
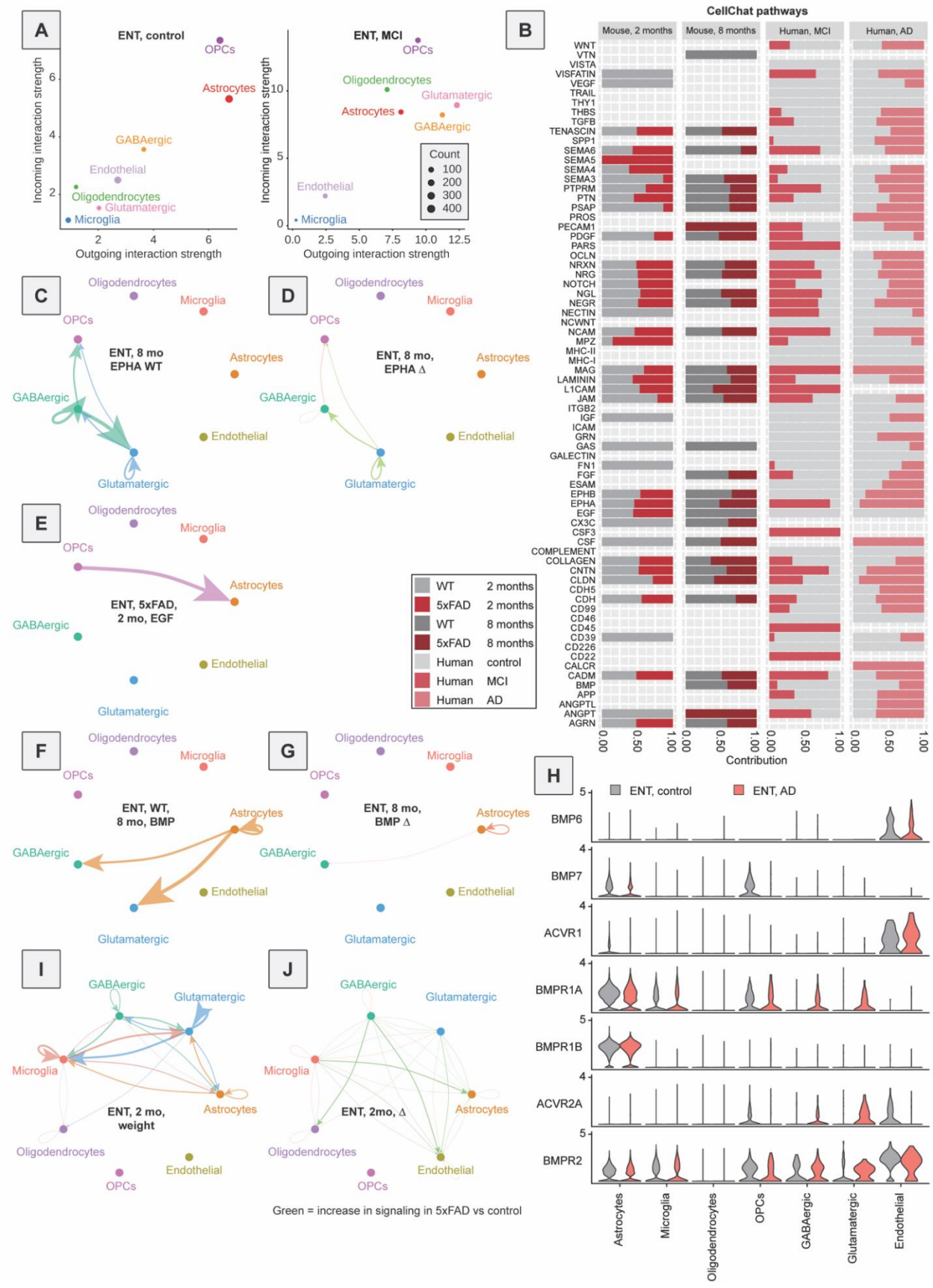

A
Micro-M1
Micro-M7
Micro-M4
Micro-M5
Micro-M2
Micro-M3
Micro-M6
Zsummary.qual
Module Size
Micro-M6
Micro-M1
Micro-M2
Micro-M5
Micro-M7
Micro-M3
Micro-M4
Zsummary.pres
Module Size
B
Astro-M1
Astro-M2
Zsummary.qual
Module Size
Astro-M2
Astro-M1
Zsummary.pres
Module Size
C
Oligo-M1
Oligo-M2
Oligo-M3
Zsummary.qual
Module Size
Oligo-M3
Oligo-M1
Oligo-M2
Zsummary.pres
Module Size
D
5xFAD vs. WT, 2 months
Oligo-M1
Oligo-M2
Oligo-M3
Average log2 (Fold change)
E
5xFAD vs. WT, 8 months
Oligo-M2
Oligo-M1
Oligo-M3
Average log2 (Fold change)
F
MCI vs. control
Oligo-M2
Oligo-M3
Average log2 (Fold change)
G
AD vs. control
Oligo-M3
Oligo-M2
Average log2 (Fold change)

Figure S3.5: Additional hdWGCNA results for GABAergic neurons, OPCs, and glutamatergic neurons.
(A) zSummary metrics indicating module quality and preservation across species for modules found within the GABAergic neuron cluster using hdWGCNA.
(B) DME analysis of GABAergic neuron modules in 2-month-old 5xFAD mice.
(C) DME analysis of GABAergic neuron modules in 8-month-old 5xFAD mice.
(D) DME analysis of GABAergic neuron modules in MCI.
(E) DME analysis of GABAergic neuron modules in AD.
(F) zSummary metrics indicating module quality and preservation across species for modules found within the OPC cluster using hdWGCNA.
(G) DME analysis of OPC modules in 2-month-old 5xFAD mice.
(H) DME analysis of OPC modules in MCI.
(I) zSummary metrics indicating module quality and preservation across species for modules found within the glutamatergic neuron cluster using hdWGCNA.
(J) DME analysis of glutamatergic neuron modules in 2-month-old 5xFAD mice.
(K) DME analysis of glutamatergic neuron modules in MCI.
(L) GO analysis of genes in the GABA-M3 hdWGCNA module using the molecular function database.


A
Zsummary.qual
Zsummary.pres
Module Size
GABA-M3
GABA-M5
GABA-M2
GABA-M7
GABA-M6
GABA-M4
GABA-M1
B
5xFAD vs. WT, 2 months
C
5xFAD vs. WT, 8 months
D
MCI vs. control
E
AD vs. control
Average log2 (Fold change)
F
OPC-M4
OPC-M5
OPC-M2
OPC-M1
OPC-M3
G
5xFAD vs. WT, 2 months
H
MCI vs. control
I
Glutamate-M1
Glutamate-M3
Glutamate-M5
Glutamate-M6
Glutamate-M4
Glutamate-M2
J
5xFAD vs. WT, 2 months
K
MCI vs. control
L
Molecular Function
P-type calcium transporter activity
ionotropic glutamate receptor binding
calcium-activated potassium channel activity
P-type ion transporter activity
type 1 fibroblast growth factor receptor binding
inorganic phosphate transmembrane transporter activity
calmodulin-dependent protein phosphatase activity
Enrichment log(combined score)

Figure S3.6: Additional CellChat and NeuronChat results.
(A) Predicted outgoing and incoming cellular signaling with CellChat for each cell type in the ENT of individuals without cognitive impairment (left) and predicted outgoing and incoming cellular signaling with CellChat for each cell type in the ENT of samples from donors with MCI.
(B) Comparison of predicted signaling for all CellChat pathways in 5xFAD versus control mice at 2 and 8 months of age (left), and in individuals with MCI or AD versus those without cognitive impairment (right).
(C) CellChat signaling inferred for EPHA in control mice at 8 months of age.
(D) Differences in CellChat signaling inferred for EPHA in 5xFAD versus control mice at 8 months of age. An increase in signaling in 5xFAD vs control is indicated in green, and a decrease in signaling in red.
(E) CellChat signaling inferred for EGF in 5xFAD mice at 2 months of age.
(F) CellChat signaling inferred for BMP in control mice at 8 months of age.
(G) Differences in CellChat signaling inferred for BMP in 5xFAD versus control mice at 8 months of age. Decreases in signaling are indicated in red.
(H) Violin plots showing expression of genes within the BMP signaling pathway in postmortem human tissue samples from patients with AD and those without cognitive impairment. Included genes are BMP ligands, BMP receptors, and Activin receptors.
(I) NeuronChat signaling inferred for all signaling pathways aggregated in control mice at 2 months of age.
(J) Differences in NeuronChat signaling inferred for all signaling pathways in 5xFAD versus control mice at 2 months of age. An increase in signaling in 5xFAD vs control is indicated in green and a decrease in signaling in red.


A
ENT, control
ENT, MCI
Incoming interaction strength
Outgoing interaction strength
OPCs
Astrocytes
GABAergic
Endothelial
Oligodendrocytes
Glutamatergic
Microglia
Count
100
200
300
400
B
CellChat pathways
Mouse, 2 months
Mouse, 8 months
Human, MCI
Human, AD
Contribution
C
ENT, 8 mo
EPHA WT
D
ENT, 8 mo,
EPHA Δ
E
ENT, 5xFAD,
2 mo, EGF
WT 2 months
5xFAD 2 months
WT 8 months
5xFAD 8 months
Human control
Human MCI
Human AD
F
ENT, WT,
8 mo, BMP
G
ENT, 8 mo,
BMP Δ
H
ENT, control
ENT, AD
BMP6
BMP7
ACVR1
BMPR1A
BMPR1B
ACVR2A
BMPR2
I
ENT, 2 mo,
weight
J
ENT, 2mo, Δ
Green = increase in signaling in 5xFAD vs control

3.6 References

1. Galton, C. J., Patterson, K., Xuereb, J. H. & Hodges, J. R. *Atypical and Typical Presentations of Alzheimer's Disease: A Clinical, Neuropsychological, Neuroimaging and Pathological Study of 13 Cases*. *Brain* vol. 123 (2000).
2. Busche, M. A. & Konnerth, A. Impairments of neural circuit function in Alzheimer's disease. *Philosophical Transactions of the Royal Society B: Biological Sciences* vol. 371 Preprint at https://doi.org/10.1098/rstb.2015.0429 (2016).
3. Busche, M. A. & Hyman, B. T. Synergy between amyloid-β and tau in Alzheimer's disease. *Nature Neuroscience* vol. 23 1183–1193 Preprint at https://doi.org/10.1038/s41593-020-0687-6 (2020).
4. De Strooper, B. & Karran, E. The Cellular Phase of Alzheimer's Disease. *Cell* vol. 164 603–615 Preprint at https://doi.org/10.1016/j.cell.2015.12.056 (2016).
5. Grubman, A. *et al.* A single-cell atlas of entorhinal cortex from individuals with Alzheimer's disease reveals cell-type-specific gene expression regulation. *Nat Neurosci* **22**, 2087–2097 (2019).
6. Mathys, H. *et al.* Single-cell transcriptomic analysis of Alzheimer's disease. *Nature* **570**, 332–337 (2019).
7. Chen, S. *et al.* Spatially resolved transcriptomics reveals genes associated with the vulnerability of middle temporal gyrus in Alzheimer's disease. *Acta Neuropathol Commun* **10**, (2022).
8. Gabitto[1], M. I. *et al.* Integrated multimodal cell atlas of Alzheimer's disease. *bioRxiv* (2024) doi:10.1101/2023.05.08.539485.
9. Efremova, M., Vento-Tormo, M., Teichmann, S. A. & Vento-Tormo, R. CellPhoneDB: inferring cell–cell communication from combined expression of multi-subunit ligand–receptor complexes. *Nat Protoc* **15**, 1484–1506 (2020).
10. Jin, S. *et al.* Inference and analysis of cell-cell communication using CellChat. *Nat Commun* **12**, (2021).
11. Chen, Y. *et al.* Exercise Modifies the Transcriptional Regulatory Features of Monocytes in Alzheimer's Patients: A Multi-Omics Integration Analysis Based on Single Cell Technology. *Front Aging Neurosci* **14**, (2022).
12. Kiani Shabestari, S. *et al.* Absence of microglia promotes diverse pathologies and early lethality in Alzheimer's disease mice. *Cell Rep* **39**, (2022).
13. Zhao, S., Ye, B., Chi, H., Cheng, C. & Liu, J. Identification of peripheral blood immune infiltration signatures and construction of monocyte-associated signatures in ovarian cancer and Alzheimer's disease using single-cell sequencing. *Heliyon* **9**, (2023).
14. Li, H. *et al.* The role of signaling crosstalk of microglia in hippocampus on progression of ageing and Alzheimer's disease. *J Pharm Anal* **13**, 788–805 (2023).
15. Dang, Y. *et al.* FTH1- and SAT1-Induced Astrocytic Ferroptosis Is Involved in Alzheimer's Disease: Evidence from Single-Cell Transcriptomic Analysis. *Pharmaceuticals* **15**, (2022).
16. Zhao, W., Johnston, K. G., Ren, H., Xu, X. & Nie, Q. Inferring neuron-neuron communications from single-cell transcriptomics through NeuronChat. *Nat Commun* **14**, (2023).
17. Igarashi, K. M. Entorhinal cortex dysfunction in Alzheimer's disease. *Trends in Neurosciences* vol. 46 124–136 Preprint at https://doi.org/10.1016/j.tins.2022.11.006 (2023).

18. Geula, C. *Abnormalities of Neural Circuitry in Alzheimer's Disease Hippocampus and Cortical Cholinergic Innervation*. https://www.neurology.org (1998).
19. van Hoesen, G. W., Hyman, B. T. & Damasio, A. R. Entorhinal cortex pathology in Alzheimer's disease. *Hippocampus* vol. 1 1–8 Preprint at https://doi.org/10.1002/hipo.450010102 (1991).
20. Di Paola, M. *et al.* Episodic memory impairment in patients with Alzheimer's disease is correlated with entorhinal cortex atrophy: A voxel-based morphometry study. *J Neurol* **254**, 774–781 (2007).
21. Forner, S. *et al.* Systematic phenotyping and characterization of the 5xFAD mouse model of Alzheimer's disease. *Sci Data* **8**, (2021).
22. Oakley, H. *et al.* Intraneuronal β-amyloid aggregates, neurodegeneration, and neuron loss in transgenic mice with five familial Alzheimer's disease mutations: Potential factors in amyloid plaque formation. *Journal of Neuroscience* **26**, 10129–10140 (2006).
23. Ohno, M. *et al.* Temporal memory deficits in Alzheimer ' s mouse models : rescue by genetic deletion of BACE1. *European Journal of Neuroscience* **23**, 251–260 (2006).
24. Stuart, T., Srivastava, A., Madad, S., Lareau, C. A. & Satija, R. Single-cell chromatin state analysis with Signac. *Nat Methods* **18**, 1333–1341 (2021).
25. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587.e29 (2021).
26. Smith, A. M. *et al.* Diverse human astrocyte and microglial transcriptional responses to Alzheimer's pathology. *Acta Neuropathol* **143**, 75–91 (2022).
27. Hafemeister, C. & Satija, R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol* **20**, (2019).
28. Ahlmann-Eltze, C. & Huber, W. glmGamPoi: Fitting Gamma-Poisson generalized linear models on single cell count data. *Bioinformatics* **36**, 5701–5702 (2020).
29. Yao, Z. *et al.* A taxonomy of transcriptomic cell types across the isocortex and hippocampal formation. *Cell* **184**, 3222-3241.e26 (2021).
30. Hodge, R. D. *et al.* Conserved cell types with divergent features in human versus mouse cortex. *Nature* **573**, 61–68 (2019).
31. Bennett, M. L. *et al.* New tools for studying microglia in the mouse and human CNS. *Proc Natl Acad Sci U S A* **113**, E1738–E1746 (2016).
32. McKenzie, A. T. *et al.* Brain Cell Type Specific Gene Expression and Co-expression Network Architectures. *Sci Rep* **8**, (2018).
33. Keirstead, H., Levine, J. & Blakemore, W. Response of the oligodendrocyte progenitor cell population (defined by NG2 labelling) to demyelination of the adult spinal cord. *Glia* **22**, 161–170 (1998).
34. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, (2008).
35. Korsunsky, I., Nathan, A., Millard, N. & Raychaudhuri, S. Presto scales Wilcoxon and auROC analyses to millions of observations. *bioRxiv* (2019) doi:10.1101/653253.
36. Morabito, S., Reese, F., Rahimzadeh, N., Miyoshi, E. & Swarup, V. hdWGCNA identifies co-expression networks in high-dimensional transcriptomics data. *Cell Reports Methods* **3**, (2023).
37. Langfelder, P. & Horvath, S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, (2008).
38. Chen, E. Y. *et al. Enrichr: Interactive and Collaborative HTML5 Gene List Enrichment Analysis Tool*. http://amp.pharm.mssm.edu/Enrichr. (2013).

39. Kuleshov, M. V. *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* **44**, W90–W97 (2016).
40. Aleksander, S. A. *et al.* The Gene Ontology knowledgebase in 2023. *Genetics* **224**, (2023).
41. Ashburner, M. *et al. Gene Ontology: Tool for the Unification of Biology The Gene Ontology Consortium**. http://www.flybase.bio.indiana.edu (2000).
42. Zhang, L., Zhang, J. & Nie, Q. DIRECT-NET: An efficient method to discover cis-regulatory elements and construct regulatory networks from single-cell multiomics data. *Sci. Adv* **8**, 7393 (2022).
43. Korsunsky, I., Nathan, A., Millard, N. & Raychaudhuri, S. Presto scales Wilcoxon and auROC analyses to millions of observations. *bioRxiv* (2019).
44. Hagberg hagberg, A. A., -Los, lanlgov, Schult, D. A. & Swart swart, P. J. *Exploring Network Structure, Dynamics, and Function Using NetworkX*. http://conference.scipy.org/proceedings/SciPy2008/paper_2 (2008).
45. Ferretti, M. T. *et al.* Sex differences in Alzheimer disease — The gateway to precision medicine. *Nature Reviews Neurology* vol. 14 457–469 Preprint at https://doi.org/10.1038/s41582-018-0032-9 (2018).
46. Sadleir, K. R., Eimer, W. A., Cole, S. L. & Vassar, R. Aβ reduction in BACE1 heterozygous null 5XFAD mice is associated with transgenic APP level. *Mol Neurodegener* **10**, (2015).
47. O'Banion, M., Coleman, P. & Callahan, L. Regional neuronal loss in aging and Alzheimer's disease: a brief review. *Seminars in the Neurosciences* **6**, 307–314 (1994).
48. Seshadri, S. *et al.* Lifetime risk of dementia and Alzheimer's disease The impact of mortality on risk estimates in the Framingham Study. (1997).
49. Habib, N. *et al.* Disease-associated astrocytes in Alzheimer's disease and aging. *Nat Neurosci* **23**, 701–706 (2020).
50. Xu, J. *et al.* Multimodal single-cell/nucleus RNA sequencing data analysis uncovers molecular networks between disease-associated microglia and astrocytes with implications for drug repurposing in Alzheimer's disease. *Genome Res* **31**, 1900–1912 (2021).
51. Habib, N. *et al.* Disease-associated astrocytes in Alzheimer's disease and aging. *Nat Neurosci* **23**, 701–706 (2020).
52. Sadick, J. S. *et al.* Astrocytes and oligodendrocytes undergo subtype-specific transcriptional changes in Alzheimer's disease. *Neuron* **110**, 1788-1805.e10 (2022).
53. Langfelder, P., Luo, R., Oldham, M. C. & Horvath, S. Is my network module preserved and reproducible? *PLoS Comput Biol* **7**, (2011).
54. Mishra, P. *et al.* Rescue of Alzheimer's disease phenotype in a mouse model by transplantation of wild-type hematopoietic stem and progenitor cells. *Cell Rep* **42**, (2023).
55. Scoyni, F. *et al.* Alzheimer's disease-induced phagocytic microglia express a specific profile of coding and non-coding RNAs. *Alzheimer's and Dementia* (2023) doi:10.1002/alz.13502.
56. Forner, S. *et al.* Systematic phenotyping and characterization of the 5xFAD mouse model of Alzheimer's disease. *Sci Data* **8**, (2021).
57. Habib, N. *et al.* Disease-associated astrocytes in Alzheimer's disease and aging. *Nat Neurosci* **23**, 701–706 (2020).
58. Raft, M. C., Miller, R. H. & Noble, M. *A Glial Progenitor Cell That Develops in Vitro into an Astrocyte or an Oligodendrocyte Depending on Culture Medium*. (1983).
59. Kondo, T. & Raff, M. Oligodendrocyte Precursor Cells Reprogrammed to Become Multipotential CNS Stem Cells. *Science (1979)* **289**, 1754–1757 (2000).

60. Rivers, L. E. *et al*. PDGFRA/NG2 glia generate myelinating oligodendrocytes and piriform projection neurons in adult mice. *Nat Neurosci* **11**, 1392–1401 (2008).
61. Zhu, X., Bergles, D. E. & Nishiyama, A. NG2 cells generate both oligodendrocytes and gray matter astrocytes. *Development* **135**, 145–157 (2008).
62. Zou, P., Wu, C., Liu, T. C. Y., Duan, R. & Yang, L. Oligodendrocyte progenitor cells in Alzheimer's disease: from physiology to pathology. *Translational Neurodegeneration* vol. 12 Preprint at https://doi.org/10.1186/s40035-023-00385-7 (2023).
63. Buchanan, J. A. *et al*. Oligodendrocyte precursor cells ingest axons in the mouse neocortex. *Proc Natl Acad Sci U S A* **119**, (2022).
64. Auguste, Y. S. S. *et al*. Oligodendrocyte precursor cells engulf synapses during circuit remodeling in mice. *Nat Neurosci* **25**, 1273–1278 (2022).
65. Akay, L. A., Effenberger, A. H. & Tsai, L.-H. Cell of all trades: oligodendrocyte precursor cells in synaptic, vascular, and immune function. (2021) doi:10.1101/gad.344218.
66. Shahoha, M., Cohen, R., Ben-Simon, Y. & Ashery, U. cAMP-Dependent Synaptic Plasticity at the Hippocampal Mossy Fiber Terminal. *Frontiers in Synaptic Neuroscience* vol. 14 Preprint at https://doi.org/10.3389/fnsyn.2022.861215 (2022).
67. Wong, S. T. *et al*. *Calcium-Stimulated Adenylyl Cyclase Activity Is Critical for Hippocampus-Dependent Long-Term Memory and Late Phase LTP*. *Neuron* vol. 23 (1999).
68. Fieblinger, T., Perez-Alvarez, A., Lamothe-Molina, P. J., Gee, C. E. & Oertner, T. G. Presynaptic cGMP sets synaptic strength in the striatum and is important for motor learning . *EMBO Rep* **23**, (2022).
69. Ricciarelli, R. *et al*. A Novel Mechanism for Cyclic Adenosine Monophosphate-Mediated Memory Formation: Role of Amyloid Beta. doi:10.1002/ana.
70. Canepa, E. *et al*. Cyclic adenosine monophosphate as an endogenous modulator of the amyloid-β precursor protein metabolism. *IUBMB Life* **65**, 127–133 (2013).
71. Calcagno, E. *et al*. Investigating the amyloid-beta enhancing effect of cGMP in neuro2a cells. *Mech Ageing Dev* **166**, 1–5 (2017).
72. Gentile, J. E., Carrizales, M. G. & Koleske, A. J. Control of Synapse Structure and Function by Actin and Its Regulators. *Cells* vol. 11 Preprint at https://doi.org/10.3390/cells11040603 (2022).
73. Gerber, K. J., Squires, K. E. & Hepler, J. R. Roles for regulator of g protein signaling proteins in synaptic signaling and plasticity. *Molecular Pharmacology* vol. 89 273–286 Preprint at https://doi.org/10.1124/mol.115.102210 (2016).
74. Catterall, W. A. & Few, A. P. Calcium Channel Regulation and Presynaptic Plasticity. *Neuron* vol. 59 882–901 Preprint at https://doi.org/10.1016/j.neuron.2008.09.005 (2008).
75. Jin, S. *et al*. Inference and analysis of cell-cell communication using CellChat. *Nat Commun* **12**, (2021).
76. Poppe, L. *et al*. EphA4 loss improves social memory performance and alters dendritic spine morphology without changes in amyloid pathology in a mouse model of Alzheimer's disease. *Alzheimers Res Ther* **11**, (2019).
77. Fu, A. K. Y. *et al*. Blockade of EphA4 signaling ameliorates hippocampal synaptic dysfunctions in mouse models of Alzheimer's disease. *Proc Natl Acad Sci U S A* **111**, 9959–9964 (2014).
78. Tamura, K., Chiu, Y.-W., Shiohara, A., Hori, Y. & Tomita, T. EphA4 regulates Aβ production via BACE1 expression in neurons. *The FASEB Journal* **34**, 16383–16396 (2020).

79. Baudet, S., Bécret, J. & Nicol, X. Approaches to manipulate ephrin-a:Epha forward signaling pathway. *Pharmaceuticals* **13**, 1–25 (2020).
80. Sobue, A. *et al.* Microglial gene signature reveals loss of homeostatic microglia associated with neurodegeneration of Alzheimer's disease. *Acta Neuropathol Commun* **9**, (2021).
81. Saito, T. *et al.* Single App knock-in mouse models of Alzheimer's disease. *Nat Neurosci* **17**, 661–663 (2014).
82. Chen, W. T. *et al.* Spatial Transcriptomics and In Situ Sequencing to Study Alzheimer's Disease. *Cell* **182**, 976-991.e19 (2020).
83. Thomas, R. *et al.* Epidermal growth factor prevents APOE4 and amyloid-beta-induced cognitive and cerebrovascular deficits in female mice. *Acta Neuropathol Commun* **4**, 111 (2016).
84. Crews, L. *et al.* Increased BMP6 levels in the brains of Alzheimer's disease patients and APP transgenic mice are accompanied by impaired neurogenesis. *Journal of Neuroscience* **30**, 12252–12262 (2010).
85. Sun, L. *et al.* Protective effects of bone morphogenetic protein 7 against amyloid-beta induced neurotoxicity in PC12 cells. *Neuroscience* **184**, 151–163 (2011).
86. Mo, X. *et al.* PTPRM Is Critical for Synapse Formation Regulated by Zinc Ion. *Front Mol Neurosci* **15**, (2022).
87. Cao, X., Dong, Y., Xing, L. & Zhang, X. A Genome-Wide Association Study of Dementia Using the Electronic Medical Record. *BioMedInformatics* **3**, 141–149 (2023).
88. Zhou, X., Sullivan, P. M., Sun, L. & Hu, F. The interaction between progranulin and prosaposin is mediated by granulins and the linker region between saposin B and C. *J Neurochem* **143**, 236–243 (2017).
89. Andersson, A. *et al.* Development of parallel reaction monitoring assays for cerebrospinal fluid proteins associated with Alzheimer's disease. *Clinica Chimica Acta* **494**, 79–93 (2019).
90. Callizot, N. *et al.* AZP2006, a new promising treatment for Alzheimer's and related diseases. *Sci Rep* **11**, (2021).
91. Masliah, E., Mallory, M., Alford, M., Deteresa, R. & Saitoh, T. *PDGF Is Associated With Neuronal and Glial Alterations of Alzheimer's Disease*. vol. 16 (1995).
92. Quaegebeur, A., Segura, I. & Carmeliet, P. Pericytes: Blood-Brain Barrier Safeguards against Neurodegeneration? *Neuron* vol. 68 321–323 Preprint at https://doi.org/10.1016/j.neuron.2010.10.024 (2010).
93. Zhao, W., Johnston, K. G., Ren, H., Xu, X. & Nie, Q. Inferring neuron-neuron communications from single-cell transcriptomics through NeuronChat. *Nat Commun* **14**, (2023).
94. Maragos, W. F., Greenamyre, J. T., Penney, J. B. & Young, A. B. *Glutamate Dysfunction in Alzheimer's Disease: An Hypothesis*. (1987).
95. Conway, M. E. Alzheimer's disease: targeting the glutamatergic system. *Biogerontology* vol. 21 257–274 Preprint at https://doi.org/10.1007/s10522-020-09860-4 (2020).
96. Hynd, M. R., Scott, H. L. & Dodd, P. R. Glutamate-mediated excitotoxicity and neurodegeneration in Alzheimer's disease. *Neurochemistry International* vol. 45 583–595 Preprint at https://doi.org/10.1016/j.neuint.2004.03.007 (2004).
97. Bhat, S. *et al.* CACNA1C (Ca v1.2) in the pathophysiology of psychiatric disease. *Progress in Neurobiology* vol. 99 1–14 Preprint at https://doi.org/10.1016/j.pneurobio.2012.06.001 (2012).
98. Bocharova, A. *et al.* Association and gene–gene interactions study of late-onset alzheimer's disease in the Russian population. *Genes (Basel)* **12**, (2021).

99. Zanos, P. *et al.* Sex-dependent modulation of age-related cognitive decline by the L-type calcium channel gene Cacna1c (Cav1.2). *European Journal of Neuroscience* **42**, 2499–2507 (2015).
100. Daschil, N. *et al.* CaV1.2 calcium channel expression in reactive astrocytes is associated with the formation of amyloid-β plaques in an Alzheimer's disease mouse model. *Journal of Alzheimer's Disease* **37**, 439–451 (2013).
101. Song, J.-Y., Ichtchenko, K., Su¨dhofsu¨dhof, T. C. & Brose, N. *Neuroligin 1 Is a Postsynaptic Cell-Adhesion Molecule of Excitatory Synapses*. vol. 96 www.pnas.org. (1999).
102. Chubykin, A. A. *et al.* Activity-Dependent Validation of Excitatory versus Inhibitory Synapses by Neuroligin-1 versus Neuroligin-2. *Neuron* **54**, 919–931 (2007).
103. Zheng, J. J. *et al.* Low expression of aging-related NRXN3 is associated with Alzheimer disease: A systematic review and meta-analysis. *Medicine (United States)* **97**, (2018).
104. Aoto, J., Földy, C., Ilcus, S. M. C., Tabuchi, K. & Südhof, T. C. Distinct circuit-dependent functions of presynaptic neurexin-3 at GABAergic and glutamatergic synapses. *Nat Neurosci* **18**, 997–1007 (2015).
105. Verduci, L., Tarcitano, E., Strano, S., Yarden, Y. & Blandino, G. CircRNAs: role in human diseases and potential use as biomarkers. *Cell Death and Disease* vol. 12 Preprint at https://doi.org/10.1038/s41419-021-03743-3 (2021).
106. Cheng, X. *et al.* F-box protein FBXO30 mediates retinoic acid receptor γ ubiquitination and regulates BMP signaling in neural tube defects. *Cell Death Dis* **10**, (2019).
107. Von Wittgenstein, J. *et al.* Sox11 is an Activity-Regulated Gene with Dentate-Gyrus-Specific Expression upon General Neural Activation. *Cerebral Cortex* **30**, 3731–3743 (2020).
108. Mathys, H. *et al.* Single-cell atlas reveals correlates of high cognitive function, dementia, and resilience to Alzheimer's disease pathology. *Cell* **186**, 4365-4385.e27 (2023).
109. Morabito, S. *et al.* Single-nucleus chromatin accessibility and transcriptomic characterization of Alzheimer's disease. *Nat Genet* **53**, 1143–1155 (2021).
110. Li, X. W. *et al.* SCAD-Brain: a public database of single cell RNA-seq data in human and mouse brains with Alzheimer's disease. *Front Aging Neurosci* **15**, (2023).
111. Park, H. *et al.* Single-cell RNA-sequencing identifies disease-associated oligodendrocytes in male APP NL-G-F and 5XFAD mice. *Nat Commun* **14**, (2023).
112. Brase, L. *et al.* Single-nucleus RNA-sequencing of autosomal dominant Alzheimer disease and risk variant carriers. *Nat Commun* **14**, (2023).
113. Xu, X., Zhang, C., Tang, G., Wang, N. & Feng, Y. Single-cell transcriptome profiling highlights the role of APP in blood vessels in assessing the risk of patients with proliferative diabetic retinopathy developing Alzheimer's disease. *Front Cell Dev Biol* **11**, (2023).
114. Liu, A., Fernandes, B. S., Citu, C. & Zhao, Z. Unraveling the intercellular communication disruption and key pathways in Alzheimer's disease: an integrative study of single-nucleus transcriptomes and genetic association. *Alzheimers Res Ther* **16**, (2024).
115. Cang, Z. *et al.* Screening cell–cell communication in spatial transcriptomics via collective optimal transport. *Nat Methods* **20**, 218–228 (2023).
116. Jensen, G. S., Leon-Palmer, N. E. & Townsend, K. L. Bone morphogenetic proteins (BMPs) in the central regulation of energy balance and adult neural plasticity. *Metabolism: Clinical and Experimental* vol. 123 Preprint at https://doi.org/10.1016/j.metabol.2021.154837 (2021).

117. Crews, L. *et al*. Increased BMP6 levels in the brains of Alzheimer's disease patients and APP transgenic mice are accompanied by impaired neurogenesis. *Journal of Neuroscience* **30**, 12252–12262 (2010).
118. Esquenazi, S., Monnerie, H., Kaplan, P. & Le Roux, P. BMP-7 and excess glutamate: Opposing effects on dendrite growth from cerebral cortical neurons in vitro. *Exp Neurol* **176**, 41–54 (2002).
119. Abe, K. & Saito, H. *Protective Effect of Epidermal Growth Factor on Glutamate Neurotoxicity in Cultured Cerebellar Neurons*. *Neuroscience Research* vol. 14 (1992).
120. Chen, X. & Holtzman, D. M. Emerging roles of innate and adaptive immunity in Alzheimer's disease. *Immunity* vol. 55 2236–2254 Preprint at https://doi.org/10.1016/j.immuni.2022.10.016 (2022).
121. Wu, K. M. *et al.* The role of the immune system in Alzheimer's disease. *Ageing Research Reviews* vol. 70 Preprint at https://doi.org/10.1016/j.arr.2021.101409 (2021).

CHAPTER 4: A comparative analysis of the retrosplenial cortex and the entorhinal cortex in AD and in 5xFAD mice

This chapter of the dissertation builds upon the previous two: it includes insights into cellular excitability from RABV tracing, characterized in chapter 2, and extends the work done in chapter 3, which focused solely on the ENT. Here, we identify the retrosplenial cortex (RSC) as a region of interest due to its significantly higher connectivity to the ENT in 5xFAD compared to control mice. We observe an increase in cellular activity in this brain region and proceed to compare and contrast the ENT and RSC via multi-omic analysis of snRNA and snATAC sequencing data from these regions in mice at multiple time points that reflect different stages of disease, from pre-clinical to cognitively impaired. We were also able to obtain postmortem tissue from these brain regions from patients with and without AD and MCI. We observe similar trends in gene expression

Dr. Desiree Macchia, Dr. Lingxuan Chen, and Dr. Guilian Tian performed sequencing, fiber photometry, viral tracing, and other experiments, and Pieter Derdeyn assisted in processing and analysis of non-sequencing-based data. Dr. Kevin T. Beier designed the study and assisted in data interpretation, and I performed computational analysis of sequencing data, assembled figures, and wrote this chapter.

4.1 Introduction

In AD research, the RSC has not been a common area of study compared to brain regions such as the hippocampus and the ENT, which have widely acknowledged roles in early stages of disease. The RSC is known to play a role in both memory and spatial navigation, and has notable connections to the hippocampus and the anterior thalamic nuclei, as assessed by axonal tracing studies[1]. In recent years there has been growing interest in the brain region, with multiple studies

implicating the RSC in age-related cognitive decline, including in memory formation and recall, as reviewed by Trask and Fournier[2]. Other research has linked the RSC to behaviors or biomarkers associated with AD in studies of mouse models and patients. In a mouse model of AD, β bursting activity in the RSC was found to be disrupted compared to controls, which the authors suggest could explain memory impairment in these mice[3]. In patients, tau accumulation in the medial parietal lobe was linked to both connectivity between the RSC and the hippocampus and assessed episodic memory[4].

In this work, we utilized RABV-based tracing[5] in the 5xFAD mouse model[6] to determine what inputs to the ENT, a brain region discussed and analyzed in chapters 1 and chapter 3 of this work, were significantly disrupted in the diseased mice compared to controls at three time points: 2, 4, and 8 months of age. While there has been some previous work investigating neuronal tracing in AD mouse models, these have largely focused on inputs to the hippocampus[7–10]. Here, significant input differences from the RSC to the ENT were observed in 5xFAD mice at early time points, which prompted further investigation of neurons in this brain region. In this work we perform analysis of not only data obtained from the RSC and ENT of 5xFAD and control mice over time, but also from postmortem tissue samples from patients with MCI or AD. This analysis across brain regions, time (mice) or diagnosis (human), and species showed trends in gene expression and cellular communication that suggest the RSC undergoes molecular changes early in disease relative to the ENT, and that these changes are impacting neuronal communication.

4.2 Methods

RABV-tracing experiments

RABV-tracing was performed in 5xFAD and control, C57BL/6, mice, similarly to tracing experiments done in chapter 2. Instead of injecting into the VTA, injections were done into the

ENT to assess differential inputs to this brain region. Rather than treating mice with various drugs to assess changes in connectivity after drug exposure, as done in chapter 2, we utilized two different genotypes to assess differences in inputs to the ENT as disease progressed: C57/BL6 background as controls, and 5xFAD mice as a model of AD. Otherwise, injections and surgery procedures were similar to those described in chapter 2 (see section 2.2 for details). The iDisco protocol[11] was used to clear and stain brains after dissection. Imaging was done on a light-sheet microscope, and resulting images were analyzed to quantify RABV-infected cells in input regions throughout the brain. Results were plotted via bar plots displayed in **Figure 4.1**.

Fiber photometry acquisition and processing

Cellular activity was assessed via calcium imaging done using fiber photometry. Fiber photometry experiments were performed as previously described[12], but with injections and optical fiber implantation instead being performed in our regions of interest: either the ENT or the RSC. Briefly, AAV-hSyn-FLExloxP-GCaMP6f was injected into either the RSC or ENT of Rbp4-Cre mice to restrict expression to layer 5 neurons. Implants were placed ~0.1 mm above the site of viral injection (within the ENT or RSC) and were secured with metal screws, Metabond, and Geristore dental epoxy. Mice were allowed to recover for approximately 2 weeks before experiments began. Mice were then recorded in the open field for 30 minutes per session at 2, 4, and 8 months. The same mice were used for each timepoint. Fiber photometry recordings were made using previously described equipment[13], and an experimental schematic is shown in **Figure 4.2E**. 465 nm and 405 nm excitation light were controlled using Synapse software and were used to stimulate calcium-dependent and isosbestic emission, respectively. Signaling measurements were filtered, processed, and corrected. To assess neural activity, we quantified the time spent

above threshold, which was set at 2.91 times the median absolute deviation of each day's recording, a value that equates to the 95% confidence interval for Gaussian data[14].

Sample collection and single-nucleus sequencing, mice

Multiomic data of the mouse ENT at 2 months and 8 months of age in 5xFAD and control mice had already been collected for the work done in chapter 3 of this dissertation (see section 3.2 for detailed methods). The same methods were utilized to collect complementary data in the mouse ENT at 4 months of age for both 5xFAD and controls. We reused the 2- and 8-month data from chapter 3 in this chapter's analyses. The RSC of 5xFAD and control mice at 2, 4, and 8 months of age were also dissected and prepared for sequencing in the same way (see section 3.2). On top of these multiomic datasets, additional snRNA-seq data was collected from 5xFAD mice and control mice at 2 and 4 months of age (both brain regions). These experiments were also done using 10x Genomics, following similar protocols, with all samples eventually being sent to UC Irvine's Genomics Research and Technology Hub for further processing and sequencing.

Sample collection and single-nucleus sequencing, human postmortem samples

Postmortem human brain tissue samples were obtained from UC Irvine's Alzheimer's Disease Research Center (ADRC). Each donor's age, sex, ethnicity, postmortem interval (PMI), apolipoprotein E (APOE) genotype, tangle stage, plaque stage, and diagnosis are reported in the table below.

Table 4.1. Metadata for each sample from donors of postmortem brain tissue. Diagnosis is shortened to AD, MCI, or cognitively normal (Norm).

Sample ID	Brain region	Age	Sex	Ethnicity	PMI	APOE	Tangle	Plaque	Diagnosis
20509r	RSC	57	F	White	3.17	3/3	6	C	AD
62909r	RSC	83	F	White	5.25	3/3	4	A	Norm
191806e	ENT	46	M	Hispanic/ American Indian	3	2/3	6	C	AD
251417e	ENT	89	F	White	5.95	2/3	3	None	Norm

290615e	ENT	58	M	White	3.83	3/4	6	C	AD
231216e	ENT	83	M	White	3.18	3/3	2	None	MCI
81417r	RSC	89	F	White	5.95	2/3	3	None	Norm
312818e	ENT	85	F	White	6.03	2/2	2	A	MCI
140615r	RSC	58	M	White	3.83	3/4	6	C	AD
103915r	RSC	93	F	White	4.15	3/4	5	A	MCI
31711r	RSC	86	F	White	6.17	3/3	3	C	MCI
132510r	RSC	57	F	White	3.17	3/3	6	C	AD
272510e	ENT	57	F	White	3.17	3/3	6	C	AD

We had collected snRNA-seq from the ENT of a patient with MCI in chapter 3 of this dissertation, and here we used the same protocols; briefly, relevant brain tissue was dissected from each sample, cells were sorted and prepared for sequencing before being submitted to UC Irvine's Genomics Research and Technology Hub for further processing and sequencing using the Parse Biosciences workflow (see section 3.2 for details).

Single-nucleus sequencing alignment and QC, mice

All mouse data was aligned to the mouse reference genome using Cell Ranger ARC (v2.0.1, 10X Genomics). Following alignment, all data were loaded into Seurat or Signac for further processing using R. We filtered out data from mice that we estimated to be doublets or low-quality nuclei by excluding any cell that did not meet all of the following QC metrics in the snRNA-seq portion of the data from downstream analysis: less than 20 percent of reads aligned to mitochondrial genes, fewer than 70,000 and more than 200 UMIs, and fewer than 10,000 and more than 75 unique genes detected. After filtering, there were a total of 42,666 nuclei from the mouse RSC at 2 months of age, 43,771 nuclei from the mouse RSC at 4 months of age, 14,722 nuclei from the mouse RSC at 8 months of age, 47,042 nuclei from the mouse ENT at 2 months of age, 29,716 nuclei from the mouse ENT at 4 months of age, and 18,052 nuclei from the mouse ENT at 8 months of age. Multiomic data was further filtered before analysis of snATAC-seq data using the following QC metrics, where only cells meeting these metrics were included in downstream

analysis: more than 10 percent reads in peaks, nucleosome signal score of less than 3, transcription start site enrichment of more than 2, less than 5 percent of reads being in a blacklist region, and counts in peaks of over 1,000 and under 70,000. Distributions of QC metrics are shown after filtering in **Figure S4.1**.

Single-nucleus sequencing alignment and QC, human postmortem samples

Data were aligned to the human reference genome using the Parse software split-pipe (v1.0.5.p, Parse Biosciences). Data were then loaded into Seurat for further processing and analysis in R. We filtered out data that was likely doublets or damaged cells by excluding any cell that did not meet all of the following QC metrics from downstream analysis: less than 10 percent of reads aligned to mitochondrial genes, fewer than 30,000 and more than 100 UMIs, and fewer than 10,000 and more than 100 unique genes detected. The RSC data were then merged into a single object and further normalization, dimensionality reduction, and analysis proceeded. The ENT data were merged and additional ENT data from previously published datasets were added[15,16]. Publicly available data was the same as used in chapter 3 of this dissertation and was processed and filtered using the same metrics (see section 3.2). Following merging with public data and filtering, there were a total of 14,165 nuclei from patients diagnosed with AD, 4,638 nuclei from patients diagnosed with MCI, and 7,927 from patients that appeared to be without cognitive deficits in the ENT dataset. In the RSC dataset, there were 11,552 nuclei from patients diagnosed with AD, 2,851 nuclei from patients diagnosed with MCI, and 4,298 from patients that appeared to be without cognitive deficits. Distributions of QC metrics are shown after filtering in **Figure S4.1**.

Annotation of cell types

Annotation of cell types was performed using a combination of label transfer from cortical mouse data downloaded from the Allen Brain Atlas[17,18], and manual annotation of cell types based

on known marker genes, as detailed in section 3.2. Similar cell types were found across species, disease condition or genotype, and brain regions, with GABAergic neurons, glutamatergic neurons, OPCs, oligodendrocytes, microglia, astrocytes, and endothelial cells all being consistently identified (**Figure 4.2, S4.2-S4.3**). The exception to this consistency was one small population of mural cells that was detected in the mouse RSC and not the mouse ENT or human data. We also annotated neuron subtypes in the mouse data as neurons were abundant – in the human data, there were less cells overall, so neuron subtypes were not analyzed. There were 5 inhibitory subclusters in both the mouse RSC and the mouse ENT, with inhibitory subclusters 1 in both the RSC and ENT (RSC Inh 1, ENT Inh 1) each expressing the marker gene *Sst*, RSC Inh 2 and ENT Inh 2 expressing *Pvalb*, RSC Inh 3 and ENT Inh 3 expressing *Lamp5*, RSC Inh 4 and ENT Inh 4 expressing *Sncg*, and RSC Inh 5 and ENT Inh 5 expressing *Vip* (**Figure S4.2-S4.3**). These inhibitory subclusters had similar patterns of gene expression across brain regions. Glutamatergic neuron subclusters on the other hand, had more variable expression across brain regions. There were 11 subclusters in each brain region. In the RSC, excitatory cluster 1 (RSC Ex 1), was inferred to be L5/6 near-projecting (NP) cortex (CTX) cells based on label transfer from the Allen Brain Atlas[17,18]. Following inferences from this label transfer, the rest of the excitatory clusters in the RSC were identified as follows: RSC Ex 2 is made up of L5 intratelencephalic (IT) CTX cells, RSC Ex 3 is made up of L4/5 IT CTX cells, RSC Ex 4 is made up of L2/3 IT CTX cells, RSC Ex 5 is made up of L2/3 IT parasubiculum, postsubiculum, and presubiculum (PPP) cells, RSC Ex 6 is made up of L4 retrosplenial and anterior cingulate (RSP-ACA) cells, RSC Ex 7 is made up of L6 IT CTX cells, RSC Ex 8 is made up of L6 corticothalamic (CT) CTX cells, RSC Ex 9 is made up of L6b CTX cells, RSC Ex 10 is made up of L5 pyramidal tract (PT) CTX cells, and RSC Ex 11 is made up of dentate gyrus (DG) cells. In the ENT, again based on label

transfer, ENT Ex 1 is made up of L5/6 NP CTX cells, ENT Ex 2 is made up of L5/6 IT temporal association, perirhinal, and ectorhinal (TPE)-ENT cells, ENT Ex 3 is made up of L2 IT ENT cells, ENT Ex 4 is made up of L2/3 IT CTX cells, ENT Ex 5 is made up of L2/3 IT ENT cells, ENT Ex 6 is made up of L2/3 IT retrohippocampal (RHP) cells, ENT Ex 7 is made up of L6 IT CTX cells, ENT Ex 8 is made up of L5 IT CTX cells, ENT Ex 9 is made up of L6 CT cells, ENT Ex 10 is made up of L5 PT CTX cells, and ENT Ex 11 is made up of DG cells. While there was some alignment between brain regions, excitatory neuron types did not have as much similarity between the RSC and the ENT as inhibitory neurons.

Differential gene expression analysis

For all available snRNA-seq data, from patients and the mouse model, we performed cell-type specific comparisons between diagnoses or genotypes using the MAST algorithm[19] implemented in Signac[20]. In mice this was done at each time point and among neuron subclusters as well. Comparisons were done between 5xFAD mice and controls, resulting in 23 comparisons at each time point of 2, 4, or 8 months of age. In data from postmortem tissue samples, comparisons were made between AD and non-cognitively impaired patients, and between MCI and non-cognitively impaired patients. There were only 7 comparisons per diagnosis, as neurons were not subclustered in the human data to comparatively more limited numbers of nuclei.

Following differential gene expression with MAST, to compare brain regions to each other across diagnosis (in humans) or across time (in mice), linear regression of log fold change values was performed. In mice, the linear relationship between log fold change in the RSC (of gene expression in 5xFAD mice compared to controls) compared to the log fold change in the ENT (of gene expression in 5xFAD mice compared to controls) was calculated for each of the following combinations of months: 2 month ENT versus 2 month RSC, 2 month ENT versus 4 month RSC,

4 month ENT versus 2 month RSC, 4 month ENT versus 4 month RSC, 4 month ENT versus 8 month RSC, 8 month ENT versus 4 month RSC, and 8 month ENT versus 8 month ENT. This analysis was then followed up with gene ontology (GO) analysis using enrichr[21,22] to assess what genes were commonly upregulated or commonly downregulated across time. A similar analysis was performed with the human data, though by diagnosis rather than time. The linear relationship between log fold change in the RSC (of gene expression in AD or MCI patients compared to controls) compared to the log fold change in the ENT (of gene expression in AD or MCI patients compared to controls) was calculated for each of the following combinations of diagnoses: AD in ENT versus AD in RSC, AD in ENT versus MCI in RSC, MCI in ENT versus AD in RSC, and MCI in ENT versus MCI in RSC. Results for this analysis are shown in **Figure 4.3**.

Cell-cell communication analysis

CellChat (v1.6.1)[23] and NeuronChat (v1.0.0)[24] were both used to infer cell-cell communication patterns in both mouse and human data based on their corresponding tutorials. Similarly to how they were used in chapter 3, CellChat and NeuronChat's standard workflows were run on each brain region, age, genotype or diagnosis individually. For mice, these were: ENT of 2-month-old 5xFAD mice, ENT of 2-month-old control mice, ENT of 4-month-old 5xFAD mice, ENT of 4-month-old control mice, ENT of 8-month-old 5xFAD mice, ENT of 8-month-old control mice, RSC of 2-month-old 5xFAD mice, RSC of 2-month-old control mice, RSC of 4-month-old 5xFAD mice, RSC of 4-month-old control mice, RSC of 8-month-old 5xFAD mice, and RSC of 8-month-old control mice. For humans, these were: ENT in AD, ENT in MCI, ENT in cognitively normal patients, RSC in AD, RSC in MCI, and RSC in cognitively normal patients. Comparative analyses for each brain region were then performed between 5xFAD and WT mice

at 2, 4, and 8 months of age. AD and cognitively normal patient samples, and MCI and cognitively normal patient samples were also compared for both RSC and ENT.

snATAC-seq data processing, mice

Alignment and initial peak calling were done using Cell Ranger ARC (v2.0.1, 10X Genomics). Data was loaded into Signac and all further processing and analysis was done using this package or DIRECT-NET[25]. Later, after cell types were annotated based on snRNA-seq data, cell-type specific peak calling was done with MACS2 implemented in Signac, which is based on the original MACS algorithm[26]. The same Signac workflow as used in chapter 3 was applied to the new ENT and to the RSC data (see section 3.2).

snATAC-seq data analysis, mice

DIRECT-NET[25] was utilized for two purposes: first, to link differentially expressed genes to accessible peaks using a machine learning-based method, and second, to infer a gene regulatory network based on known transcription factors that bind to regions in the differentially accessible, identified peak to gene linkages. Presto[27] was utilized to identify differentially expressed genes efficiently using a Wilcoxon Rank Sum test in glutamatergic neurons, the cell type of interest. Differential peak accessibility was assessed using a logistic regression framework[28] implemented in Signac. For differential peak accessibility, due to minimal differences in most conditions, a low threshold of a log fold change of 0.05 was set. The results in the RSC at 2 months and the ENT at 4 months resulted in hundreds of differentially accessible peaks each, and thus there were too many terms in the resulting network to visualize with this threshold. Thus, for visualization of networks in **Figure 4.6**, peaks with a log fold change accessibility of over 0.15 were used for network construction for these two conditions only. The resulting networks were still significantly larger

than other networks despite this higher threshold. Visualization, as done in chapter 3, was done in Python using the package networkx[29] (see section 3.2 for details).

4.3 Results

Identification of the RSC through RABV-based retrograde tracing

In chapter 2 we explored RABV-based tracing of neuronal connections in various conditions and our findings suggested that RABV tracing could be indicative of differences in cellular activity. In chapter 3, we performed an analysis of the ENT in AD and in 5xFAD mice, which resulted in glutamatergic signaling and related processes being identified as significant changes in disease. However, in our multiomic analysis, we did not see major differences early in pathogenesis (at 2 months of age in 5xFAD mice). Because of the importance of glutamatergic neurons in the ENT, and the lack of identifiable transcriptomic and epigenetic changes in the ENT at 2 months of age, we decided to utilize RABV-tracing to identify brain regions that may be implicated earlier in disease, thus helping to explain the later changes seen in ENT neurons. Injections of the retrograde RABV tracer into the ENT were done in 5xFAD and control mice at 2, 4, and 8 months of age (**Figure 4.1A**). Following injection, whole brains were cleared, immunostained, and imaged using the iDisco protocol[11] and light-sheet microscopy (ZEISS Z.1). These images were then used to count rabies-labeled cells per input brain region and plotted in bar graphs by genotype and age as shown in **Figure 4.1B-D**. There were differences in RABV-labeled inputs from the subiculum and the RSC, at all time points, but this was only significant at 4 months of age (**Figure 4.1C**). We chose to focus on the RSC due to the location of the subiculum being adjacent to the hippocampus and the ENT, both widely studied in the context of AD, and due to the lack of knowledge surrounding the RSC in this disease. Further study of the RSC was also

warranted based on differences in cellular activity as assessed by calcium imaging via fiber photometry experiments (**Figure 4.1E-H**). Spontaneous activity, as measured by percent of time

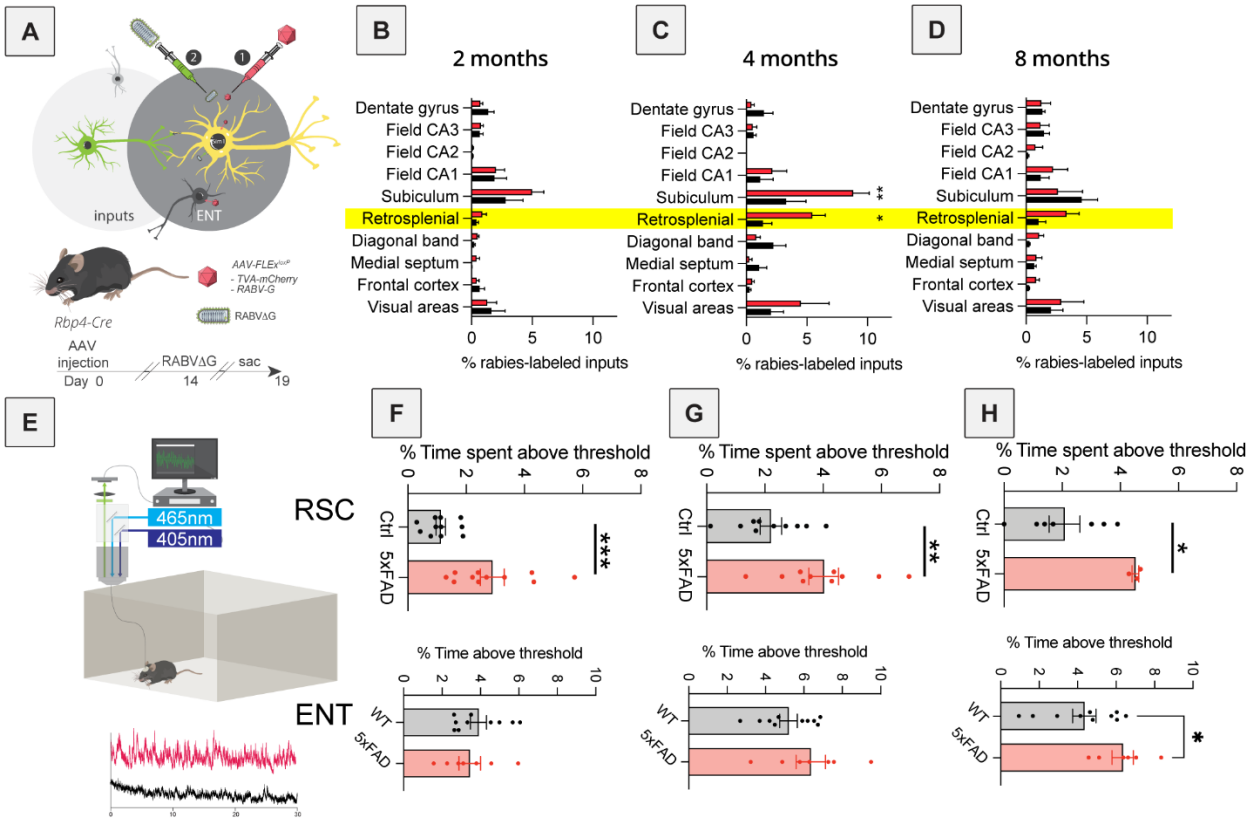


Figure 4.1: RABV-tracing and fiber photometry results in 5xFAD versus control mice.
(A) Schematic of injections
(B) Bar plot comparing rabies-labeled inputs to the ENT from 5xFAD (red) and control (black) mice at 2 months of age.
(C) Bar plot comparing rabies-labeled inputs to the ENT from 5xFAD (red) and control (black) mice at 4 months of age. Here, a significant difference is detected between 5xFAD and control mice in the subiculum and the RSC.
(D) Bar plot comparing rabies-labeled inputs to the ENT from 5xFAD (red) and control (black) mice at 8 months of age.
(E) Schematic of fiber photometry experimental set up.
(F) Percent time spent above the threshold for 5xFAD and control mice in the RSC (top) and ENT (bottom) at 2 months of age. Here a significant difference is detected between 5xFAD and control mice in the RSC and not the ENT.
(G) Percent time spent above the threshold for 5xFAD and control mice in the RSC (top) and ENT (bottom) at 4 months of age. Here a significant difference is detected between 5xFAD and control mice in the RSC and not the ENT.
(H) Percent time spent above the threshold for 5xFAD and control mice in the RSC (top) and ENT (bottom) at 8 months of age. Here a significant difference is detected between 5xFAD and control mice in the RSC and the ENT.

spent above the threshold, was significantly higher in 5xFAD mice compared to controls as early as 2-months of age in the RSC (**Figure 4.1F**), and this hyperactivity was sustained at 4- (**Figure 4.1G**) and 8- months of age (**Figure 4.1H**). Hyperactivity in the ENT of 5xFAD mice was not observed until mice were 8 months old (**Figure 4.1F-H**). This observation supports the hypothesis that RABV tracing results can be indicative of increased neuronal activity from input regions. Given that the early changes in the RSC's hyperactivity seem to reflect later changes in the ENT, we next wanted to explore what molecular changes may be underlying these brain-region specific neuronal differences.

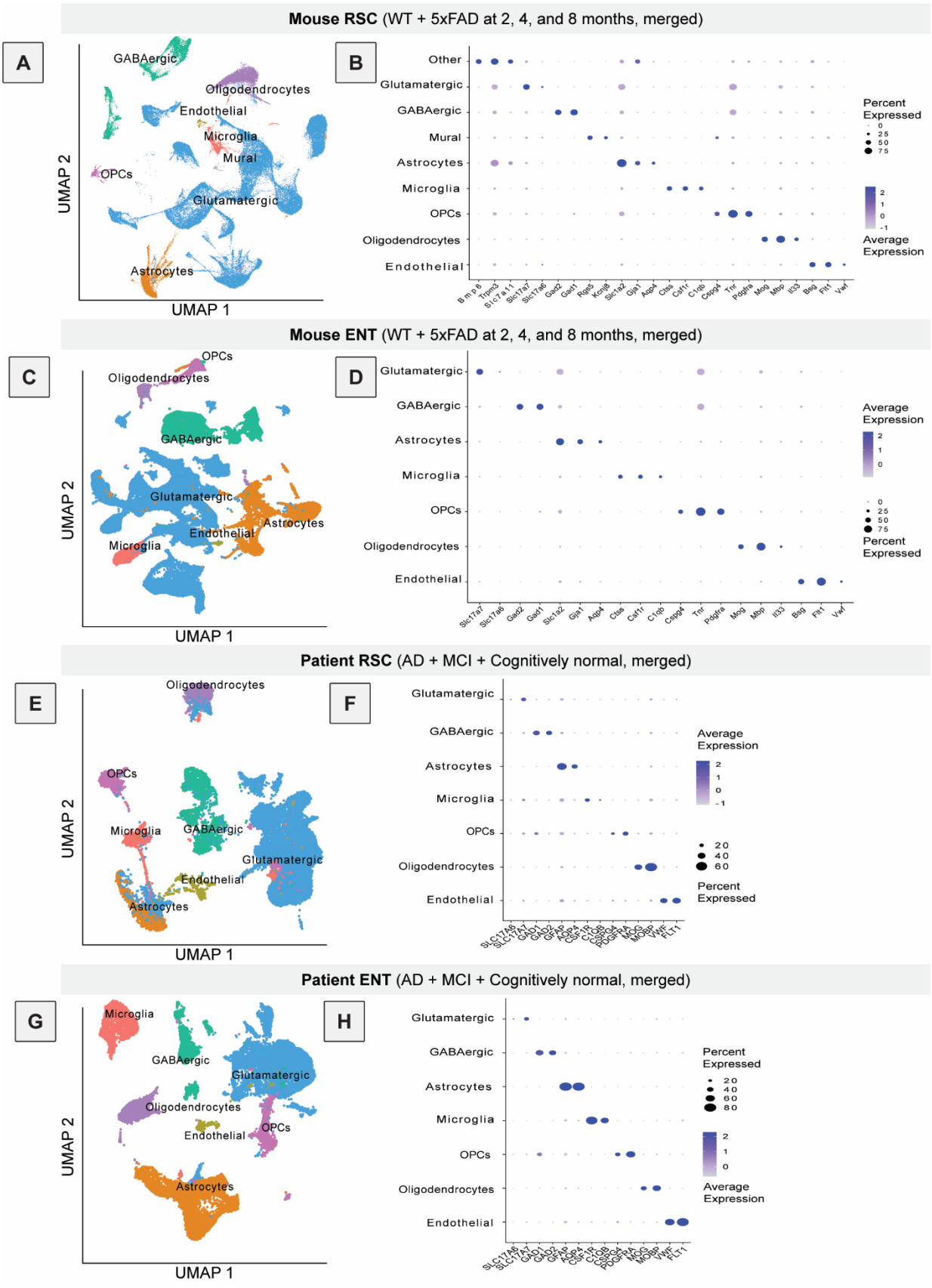
Characterization of cell types across brain regions and species

To assess differences between the mouse RSC and ENT across pathogenesis, we performed snRNA-seq and snATAC-seq at three time points: 2 months, to observe changes before cognitive deficits or major plaque accumulation has occurred, 4 months, to observe changes after plaque accumulation has been established and cognition has started to be affected, and at 8 months of age, when cognitive decline and Aβ plaques are both evident. Cell type annotation was performed based on known marker genes and their expression in the snRNA-seq data, as well as label transfer from previously published and well-annotated datasets – the resulting cell types were visualized in 2D using UMAP. For the mouse RSC, this UMAP and gene expression is shown in **Figure 4.2A-B**, and for the mouse ENT these are shown in **Figure 4.2C-D**. Seven cell types were found in both brain regions: OPCs, oligodendrocytes, astrocytes, microglia, endothelial cells, GABAergic neurons, and glutamatergic neurons. Subclustering of GABAergic neurons in the RSC yielded 5 inhibitory subtypes: RSC Inh 1, RSC Inh 2, RSC Inh 3, RSC Inh 4, and RSC Inh 5. RSC glutamatergic neurons were subclustered into 11 excitatory subtypes: RSC Ex 1, RSC Ex 2, RSC

Ex 3, RSC Ex 4, RSC Ex 5, RSC Ex 6, RSC Ex 7, RSC Ex 8, RSC Ex 9, RSC Ex 10, and RSC Ex 11. These subtypes, their gene expression, and labels as inferred by label transfer are shown in **Figure S4.2**, demonstrating distinctive gene expression patterns in each subcluster. Similar subclustering and annotation was done in the ENT, again resulting in 5 inhibitory subtypes (ENT Inh 1-5) and 11 excitatory subtypes (ENT Ex 1-11, **Figure S4.3**).

In the human data, the same 7 cell types detected in the mouse data listed above were found, with annotation again being based on marker genes that are common between species (**Figure 4.2E-H**). Subclustering of neurons was not performed due to limited numbers of neurons in the human data compared to the mouse data. The human RSC data (**Figure 4.2E-F**) and the human ENT data (**Figure 4.2G-H**) used here, unlike the ENT data included in chapter 3, came from a mix of male and female donors. Overall, the human data had a similar pattern of gene expression per cell type compared to the mouse data, and we proceeded to analyze AD and 5xFAD samples compared to controls in each brain region.

Figure 4.2: Cell type annotation and marker gene expression.
(A) Dimensionality reduction of snRNA-seq data from the mouse RSC in 5xFAD and control mice at 2, 4, and 8 months of age, with general cell type annotations displayed. Cell types are colored so that: glutamatergic are blue, GABAergic are green, oligodendrocytes are purple, OPCs are pink, microglia are red, astrocytes are orange, mural cells are grey, and endothelial cells are yellow.
(B) Dotplot showing expression of known marker genes in the annotated snRNA-seq data for the mouse RSC.
(C) Dimensionality reduction of snRNA-seq data from the mouse ENT in 5xFAD and control mice at 2, 4, and 8 months of age, with general cell type annotations displayed.
(D) Dotplot showing expression of known marker genes in the annotated snRNA-seq data for the mouse ENT.
(E) Dimensionality reduction of snRNA-seq data from the human RSC in patients with AD, with MCI, or without cognitive deficits, with general cell type annotations displayed.
(F) Dotplot showing expression of known marker genes in the annotated snRNA-seq data for the human RSC.
(G) Dimensionality reduction of snRNA-seq data from the human ENT in patients with AD, with MCI, or without cognitive deficits, with general cell type annotations displayed.
(H) Dotplot showing expression of known marker genes in the annotated snRNA-seq data for the human ENT.


Mouse RSC (WT + 5xFAD at 2, 4, and 8 months, merged)
A
GABAergic
Oligodendrocytes
Endothelial
Microglia
Mural
OPCs
Glutamatergic
Astrocytes
UMAP 2
UMAP 1
B
Other
Glutamatergic
GABAergic
Mural
Astrocytes
Microglia
OPCs
Oligodendrocytes
Endothelial
Percent Expressed
Average Expression
Mouse ENT (WT + 5xFAD at 2, 4, and 8 months, merged)
C
OPCs
Oligodendrocytes
GABAergic
Glutamatergic
Astrocytes
Endothelial
Microglia
UMAP 2
UMAP 1
D
Glutamatergic
GABAergic
Astrocytes
Microglia
OPCs
Oligodendrocytes
Endothelial
Average Expression
Percent Expressed
Patient RSC (AD + MCI + Cognitively normal, merged)
E
Oligodendrocytes
OPCs
Microglia
GABAergic
Glutamatergic
Endothelial
Astrocytes
UMAP 2
UMAP 1
F
Glutamatergic
GABAergic
Astrocytes
Microglia
OPCs
Oligodendrocytes
Endothelial
Average Expression
Percent Expressed
Patient ENT (AD + MCI + Cognitively normal, merged)
G
Microglia
GABAergic
Glutamatergic
Oligodendrocytes
Endothelial
OPCs
Astrocytes
UMAP 2
UMAP 1
H
Glutamatergic
GABAergic
Astrocytes
Microglia
OPCs
Oligodendrocytes
Endothelial
Percent Expressed
Average Expression

Gene expression differences and correlations across brain regions, time

We performed differential gene expression analysis in each cell type and at each time point in mice; 5xFAD mice were compared to controls, with a positive log fold change being indicative of upregulation in 5xFAD, and a negative log fold change being indicative of downregulation in 5xFAD compared to control mice. Following differential gene expression, we utilized linear regression to evaluate relationships between gene expression changes at different months across brain regions. The linear relationship between the log fold change values for each time point in the RSC to the log fold change values for each time point in the ENT was calculated and the resulting regression lines for each cell type were plotted, with a positive slope indicating a positive correlation between gene expression difference in the two brain regions, and a negative slope indicating generally opposing gene expression patterns (**Figure 4.3A-G**).

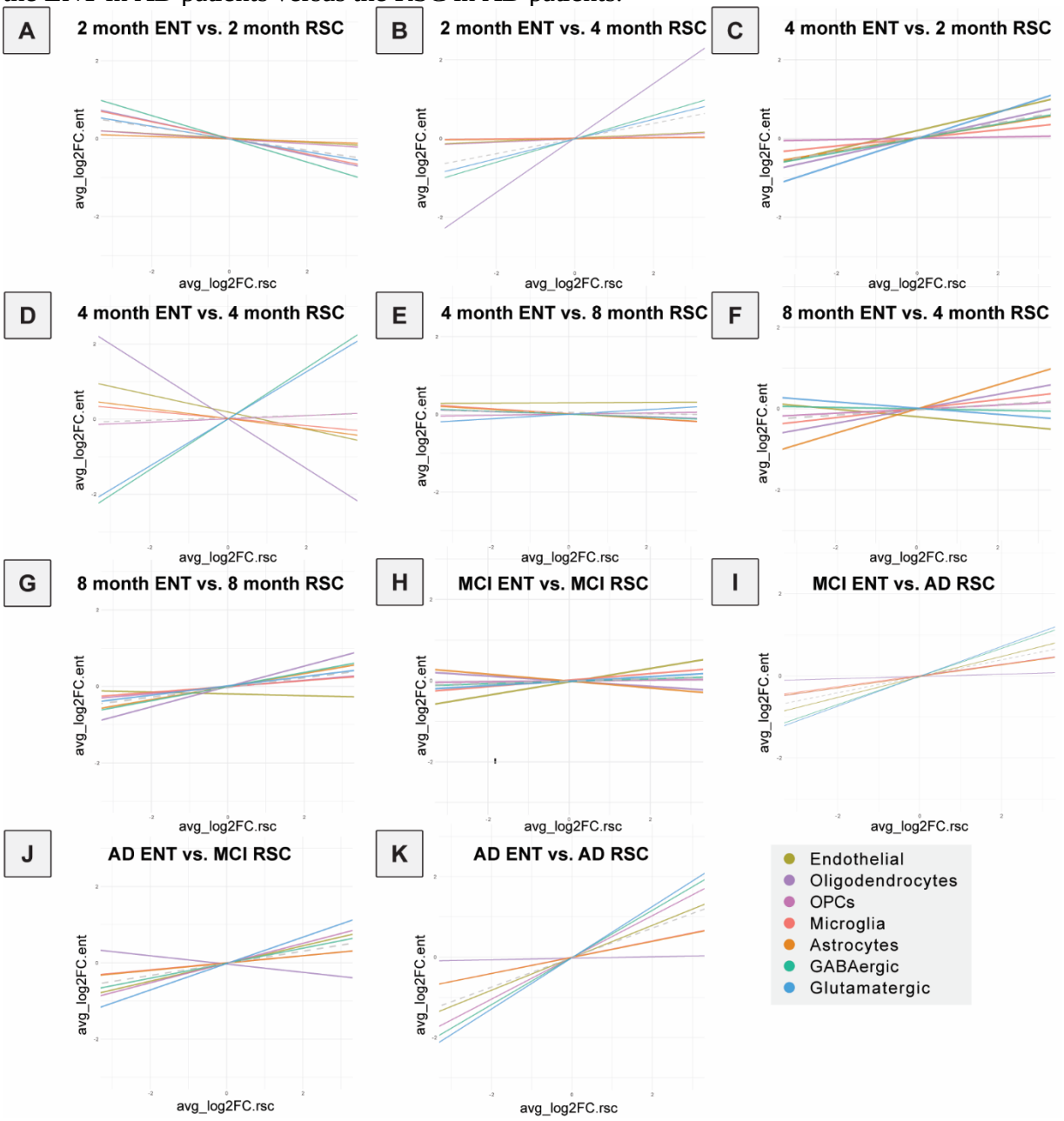
A negative slope was observed for every cell type in the 2-month ENT versus 2-month RSC comparison, indicating that gene expression between the two brain regions at this time point did not align, with many genes upregulated in the RSC being downregulated in the ENT, and vice versa (**Figure 4.3A**). This difference, which was strongest in GABAergic neurons, may reflect the differences seen in cellular activity (**Figure 4.1F**). The 4-month ENT versus 2-month RSC relationships and the 2-month ENT versus 4-month RSC relationships were all positive or close to zero (**Figure 4.3B-C**), indicating a greater level of similarity between disruptions at these different time points. In the 2-month ENT to 4-month RSC comparison, oligodendrocytes stood out as the cell type with the most positive correlation between gene expression changes (**Figure 4.3B**). In the opposite comparison (2-month RSC to 4-month ENT), glutamatergic neurons were the most positively correlated cell type between brain regions (**Figure 4.3C**). The 4-month ENT versus 4-month RSC relationships showed great variability, with neurons being highly positively correlated,

but glial cells generally being negatively correlated (**Figure 4.3D**). The 4 month versus 8-month relationships (**Figure 4.3E-F**) did not show a strong correlation either positively or negatively overall, though the 4-month RSC did appear to be slightly more predictive of 8-month ENT changes than the other way around. Finally, the 8-month ENT versus the 8-month RSC changes generally had a positive correlation to each other (**Figure 4.3G**). Compared to the negative relationships at 2 months of age and the mixed relationships at 4 months of age, this positive relationship at 8 months of age indicates that as disease progresses, changes by brain region become more similar; at early stages there are major regional differences, but these appear to fade once pathology becomes more severe.

These results suggest that at early time points, changes in oligodendrocytes in the ENT may be predictive of changes in the RSC, and changes in glutamatergic neurons in the RSC may be predictive of changes in the ENT. The latter relationship is of particular interest given the results from neuronal tracing and neuronal activity differences, but changes in glial cells are also worth exploring, especially given their importance in AD based on previous study[30–35], and interactions with and impact on neurons.

Figure 4.3: Linear regression of human AD, MCI and mouse 5xFAD log fold change in gene expression.
(A) Plot showing the linear relationship between log fold change (log2FC) in gene expression in the mouse ENT at 2 months of age versus the mouse RSC at 2 months of age. Each line is the relationship between a cell type (key in bottom right corner of **Figure 4.3**), and the dashed grey line is the average of all cell types.
(B) Plot showing the linear relationship between log fold change (log2FC) in gene expression in the mouse ENT at 2 months of age versus the mouse RSC at 4 months of age.
(C) Plot showing the linear relationship between log fold change (log2FC) in gene expression in the mouse ENT at 4 months of age versus the mouse RSC at 2 months of age.
(D) Plot showing the linear relationship between log fold change (log2FC) in gene expression in the mouse ENT at 4 months of age versus the mouse RSC at 4 months of age.
(E) Plot showing the linear relationship between log fold change (log2FC) in gene expression in the mouse ENT at 4 months of age versus the mouse RSC at 8 months of age.
(F) Plot showing the linear relationship between log fold change (log2FC) in gene expression in the mouse ENT at 8 months of age versus the mouse RSC at 4 months of age.

(G) Plot showing the linear relationship between log fold change (log2FC) in gene expression in the mouse ENT at 8 months of age versus the mouse RSC at 8 months of age.
(H) Plot showing the linear relationship between log fold change (log2FC) in gene expression in the ENT in MCI patients versus the RSC in MCI patients.
(I) Plot showing the linear relationship between log fold change (log2FC) in gene expression in the ENT in MCI patients versus the RSC in AD patients.
(J) Plot showing the linear relationship between log fold change (log2FC) in gene expression in the ENT in AD patients versus the RSC in MCI patients.
(K) Plot showing the linear relationship between log fold change (log2FC) in gene expression in the ENT in AD patients versus the RSC in AD patients.



We next repeated these analyses in the human data and observed similar trends (**Figure 4.3H-K**). When comparing gene expression changes in the RSC of patients with MCI to changes in the ENT of patients with MCI, there were mixed results (**Figure 4.3H**), with some cell types showing a positive correlation and some showing a negative correlation. This result was most similar to the 4-month versus 4-month region comparison in the mouse model. The relationships between the ENT in MCI versus the RSC in AD, and the relationships between the RSC in MCI versus the ENT in AD were largely positive, except for oligodendrocytes in the latter comparison, where this cell type had a negative correlation (**Figure 4.3I-J**). For comparisons of MCI to AD or AD to AD (**Figure 4.3I-K**), glutamatergic neurons were consistently the most highly correlated cell type from one brain region to the other, highlighting that at later stages, changes in this cell type are common to both brain regions. The overall relationship between changes in the ENT in AD versus changes in the RSC in AD (**Figure 4.3K**) was the most positive of any comparison done between brain regions for both mouse and patient data. This strong relationship shown in **Figure 4.3K** supports the hypothesis that as disease progresses, in this case from MCI to AD (in mice from 2 to 8 months), changes by brain region progressively become more similar, while at earlier time points (2- or 4-months of age in 5xFAD mice, MCI in patients), gene expression changes are distinctive between brain regions and even show an inverse relationship.

The strong positive correlation between gene expression changes in glutamatergic neurons of the ENT and RSC in AD and MCI shown in **Figure 4.3I-K**, in combination with the correlation in mice observed in **Figure 4.3C**, indicated that glutamatergic neurons in the RSC are undergoing changes early in disease progression that may reflect later stages of disease. To assess what types of genes were contributing to these changes, we performed GO analysis of differentially expressed genes (log-fold change > 0.1 or log-fold change < -0.1) in glutamatergic neurons in AD and MCI.

This was done for two groups of genes: first, genes that were upregulated in both the ENT and the RSC in AD (**Figure 4.4A-C**), to assess terms associated with this late stage of disease regardless of brain region. And second, for genes that were upregulated in the RSC in MCI and downregulated in the ENT in MCI (**Figure 4.4D-F**), to assess terms that were associated to genes that had opposing expression patterns in the RSC compared to the ENT at this earlier disease stage.

Terms that appeared in the first group, those that were up in AD, were largely associated with synapses, ion channels, and neurotransmitters including both GABA and glutamate. In the second group, there was quite a bit of overlap, with one term in particular appearing to be highly associated with both groups, (Neuron Projection, **Figure 4.4C, F**), and others, while not identical,

Figure 4.4: Gene ontology analysis results from human data and differential gene expression in glutamatergic neurons in 5xFAD mice.
(A) Gene ontology analysis based on genes that are upregulated in both the RSC in AD and the ENT in AD. The top 10 most highly associated terms for the Molecular Function (2023) database are displayed.
(B) Gene ontology analysis based on genes that are upregulated in both the RSC in AD and the ENT in AD. The top 10 most highly associated terms for the Biological Process (2023) database are displayed.
(C) Gene ontology analysis based on genes that are upregulated in both the RSC in AD and the ENT in AD. The top 10 most highly associated terms for the Cellular Component (2023) database are displayed.
(D) Gene ontology analysis based on genes that are upregulated in the RSC in MCI and downregulated in the ENT in MCI. The top 10 most highly associated terms for the Molecular Function (2023) database are displayed. *The sixth term from the top is Transmitter-Gated Monoatomic Ion Channel Activity Involved In Regulation of Postsynaptic Membrane Potential.
(E) Gene ontology analysis based on genes that are upregulated in the RSC in MCI and downregulated in the ENT in MCI. The top 10 most highly associated terms for the Biological Process (2023) database are displayed.
(F) Gene ontology analysis based on genes that are upregulated in the RSC in MCI and downregulated in the ENT in MCI. The top 10 most highly associated terms for the Cellular Component (2023) database are displayed.
(G) Volcano plot showing differential gene expression analysis results in 5xFAD versus control mice. These results are from glutamatergic neurons in the RSC at 4 months of age, and genes labeled are a subset of genes within the Postsynaptic Density (GO: 0014069) term.
(H) Volcano plot showing differential gene expression analysis results in 5xFAD versus control mice. These results are from glutamatergic neurons in the ENT at 4 months of age, and genes labeled are a subset of genes within the Postsynaptic Density (GO: 0014069) term.


A
Up in AD
Molecular Function
Adenylate Cyclase Regulator Activity
Cyclase Activator Activity
alpha-N-acetylneuraminate Alpha-2,8-Sialyltransferase Activity
Leucine Binding
Calcium-Dependent Protein Serine /Threonine Phosphatase Activity
Aspartic-Type Endopeptidase Inhibitor Activity
Ionotropic Glutamate Receptor Binding
mRNA 3'-UTR AU-rich Region Binding
Histone Acetyltransferase Binding
Glutamate Receptor Binding
Enrichment log(combined score)
B
Up in AD
Biological Process
Neuronal Ion Channel Clustering
Neurotransmitter Receptor Transport, Postsynaptic Endosome To Lysosome
Regulation Of Short-Term Neuronal Synaptic Plasticity
Neuron Projection Maintenance
Intermediate Filament Bundle Assembly
Dendritic Spine Maintenance
Negative Regulation Of Microtubule Polymerization
Neuron Projection Organization
Cyclooxygenase Pathway
Regulation Of Neuronal Synaptic Plasticity
Enrichment log(combined score)
C
Up in AD
Cellular Component
Clathrin-Sculpted Gamma-Aminobutyric Acid Transport Vesicle Membrane
Clathrin-Sculpted Gamma-Aminobutyric Acid Transport Vesicle
Asymmetric, Glutamatergic, Excitatory Synapse
Axon
Inhibitory Synapse
Neuron Projection
Axonal Growth Cone
Apical Dendrite
Glutamatergic Synapse
Catenin Complex
Enrichment log(combined score)
D
Up in MCI RSC, Down in MCI ENT
Molecular Function
Neurotransmitter Receptor Activity Involved In Regulation Of Postsynaptic Membrane Potential
Ligand-Gated Monoatomic Ion Channel Activity Involved In Regulation Of Presynaptic Membrane Potential
Ionotropic Glutamate Receptor Activity
Transmitter-Gated Monoatomic Ion Channel Activity
Ligand-Gated Monoatomic Ion Channel Activity
Transmitter-Gated Monoatomic Ion Channel Activity Involved In Regulation Of Postsynaptic ...*
Glycine Binding
Transmembrane Receptor Protein Tyrosine Phosphatase Activity
Transmembrane Receptor Protein Phosphatase Activity
Glutamate Receptor Activity
Enrichment log(combined score)
E
Up in MCI RSC, Down in MCI ENT
Biological Process
Chondrocyte Development
Regulation Of Cellular Response To Insulin Stimulus
Synaptic Transmission, Glutamatergic
Protein Localization To Cytoplasmic Stress Granule
Glutamate Receptor Signaling Pathway
Regulation Of Trans-Synaptic Signaling
Regulation Of RNA Export From Nucleus
Regulation Of Dendritic Spine Maintenance
Cell-Cell Signaling By Wnt
Neuromuscular Synaptic Transmission
Enrichment log(combined score)
F
Up in MCI RSC, Down in MCI ENT
Cellular Component
Postsynaptic Density Membrane
Postsynaptic Specialization Membrane
Ionotropic Glutamate Receptor Complex
Postsynaptic Density
Clathrin Coat Of Coated Pit
AMPA Glutamate Receptor Complex
NMDA Selective Glutamate Receptor Complex
Cytoskeleton Of Presynaptic Active Zone
Neuron Projection
Clathrin Coat Of Endocytic Vesicle
Enrichment log(combined score)
G
RSC, 4 month Glutamatergic
-log(p_val_adj)
avg_log2FC
Rps27
Rps19
Rpl12
Shisa6
Rpl14
Rps18
Inpp4a
Rps14
Dnm3
Grm5
Grin1
Gria1
Rpl8
Rps3
Kcnd3
Shisa9
Ripeb4
Grik3
Gria4
Fmr1
Adgrl1
Kcnd2
Slc8a3
Grm1
Tmem108
Rtn1
Akap5
Lrp4
H
ENT, 4 month Glutamatergic
-log(p_val_adj)
avg_log2FC
Rpl30
Rps14
Rpl14
Kcnd3
Shank1
Shank3
Nlgn1
Neto1
Pdlim5
Cacng8
Syt11
Kalrn
Atp2b2
Usp50
Homer2
Grin2a
Ptk2b
Rtn4
Cacna1c
Gphn
Grin1
Grik4
Prnp
Slc8a2
Nsf
Cpeb4
Fyn
Cdk5r1
Cdh2
Shisa6
Ptprs
Neto2
Dagla
Adgra1
Cdk5

containing overlapping groups of genes related to glutamate receptor activity and synaptic transmission or plasticity. This overlap indicated a similarity between the RSC early in disease and later in disease that reflected the neuronal activity differences we had observed in 5xFAD mice (**Figure 4.1E-H**). Given this seeming concordance, we next checked for differential expression of genes of interest in the glutamatergic neurons in the mouse data (**Figure 4.4G-H, S4.4**). Genes highlighted and labeled in **Figure 4.4G-H** and **Figure S4.4** are genes found in the postsynaptic density GO term from the Cellular Component database (**Figure 4.4F**). This term was chosen as it is broad and encompasses aspects of other terms of interest; within it are genes that overlap with glutamate signaling-associated terms and various ion channel genes – all cellular components, molecular functions, or biological processes that were seen in our human AD GO analysis.

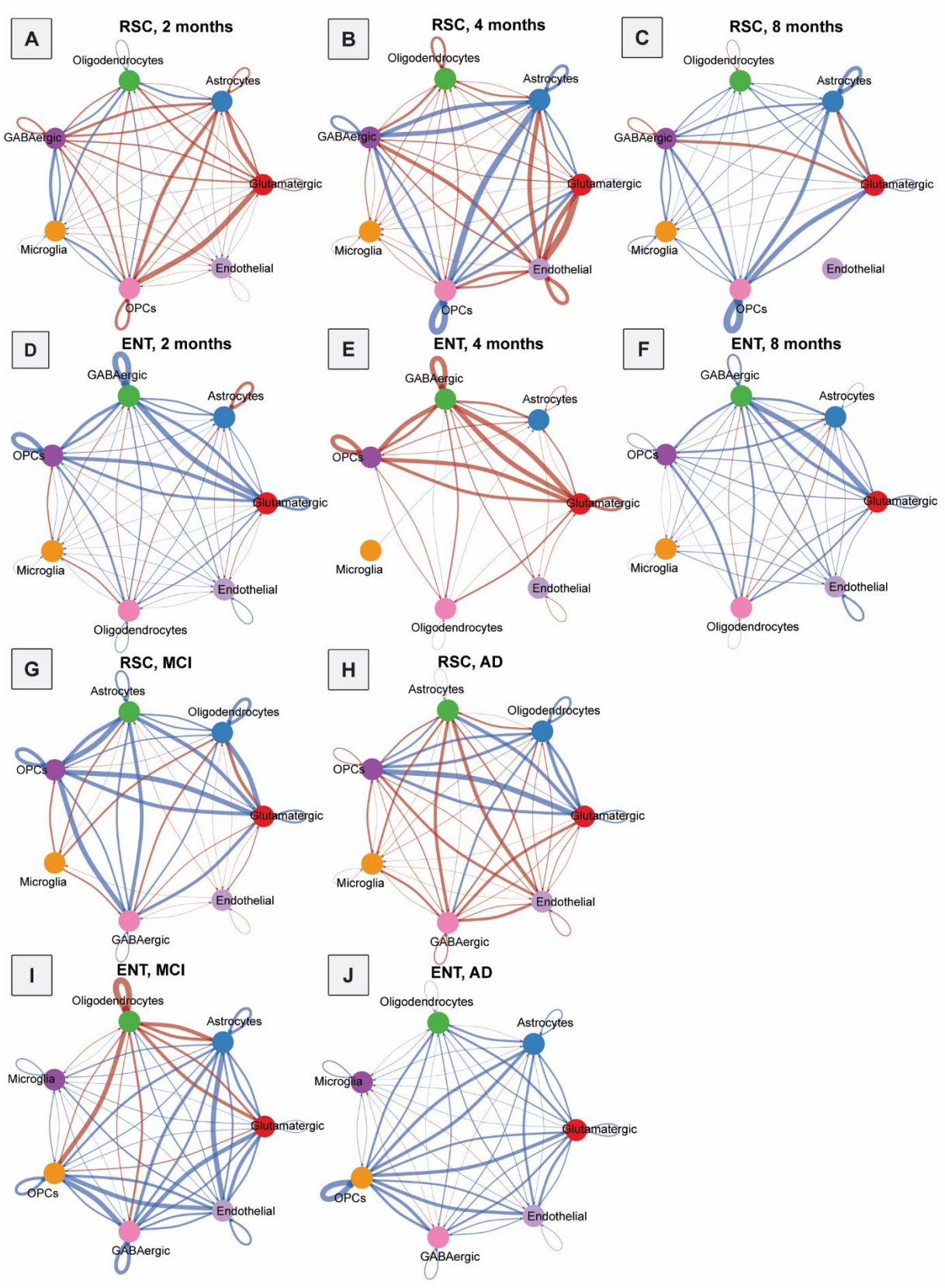
Results in mice did not reflect what we had seen in humans, with the differences at 2 months of age being the most unexpected: genes associated with postsynaptic density were more commonly upregulated in the ENT of 5xFAD mice than in the RSC. While in the ENT there were many significantly upregulated genes at this time point (**Figure S4.4B**), in the RSC results were split, with slightly more genes being downregulated than upregulated (**Figure S4.4A**). At 4-months of age, at first glance, things appeared to be more similar between the ENT and RSC: there was a relatively even split in both brain regions between up and downregulated genes (**Figure 4.4G-H**). In both the RSC and ENT glutamatergic neuron populations, the following genes were upregulated in 5xFAD mice: *Shank1*, *Fyn*, *Adcy8*, *Cpeb4*, *Cacng3*, *Arhgap32*, *Magi2*, and *Gphn*. However, taking a closer look at genes and their directionality, we saw differences including the upregulation of the following genes in the RSC in 5xFAD mice: *Dlg4*, *Adgrl1*, *Atp2b2*, *Cdk5r1*, *Syt11*, *Grin1*, *Add1*, *Srgap3*, and *Kcnd3* (all were downregulated in the ENT in 5xFAD mice). The 8-month time point was similar, in that there was a relatively even split of up and downregulated genes in both

the RSC and ENT. Notably, differential gene expression analysis at this time point yielded smaller log fold changes on average (**Figure S4.4C-D**), which may be partially explained by the smaller number of cells at this time point (**Figure S4.2A**, **S4.3A**). These expression patterns together indicate that while there may be differences in postsynaptic density in both the ENT and the RSC, the genes underlying this difference in the mouse RSC versus the mouse ENT glutamatergic neurons are distinctive, especially at 4 months of age.

Cell-cell communication changes in disease

To further explore changes over time by brain region, and to consider the system as a whole rather than individual cell types, we inferred how cellular communication between neurons and glia was altered at each time point and genotype in mice and in each diagnosis in humans, using both CellChat[23] and NeuronChat[24] (**Figure 4.5, S4.5-S4.6**). In mice, similar to the results seen in our regression analyses (**Figure 4.3A**), we observed major differences in signaling at 2 months of age (**Figure 4.5A, D**). The ENT appeared to have an overall decrease in cell-cell communication in 5xFAD mice, while the RSC had an increase. At 4 months of age, there were still apparent

Figure 4.5: Cellular communication differences inferred by CellChat across time, disease progression, in the RSC and ENT.
(A) Differences in cell-cell communication in the RSC at 2 months of age. For every panel in **Figure 4.5A-F**, nodes are colored by cell type, and edges are colored red to indicate an increase in signaling in 5xFAD compared to control mice, or blue to indicate a decrease in signaling in 5xFAD compared to control mice.
(B) Differences in cell-cell communication in the RSC at 4 months of age.
(C) Differences in cell-cell communication in the RSC at 8 months of age.
(D) Differences in cell-cell communication in the ENT at 2 months of age.
(E) Differences in cell-cell communication in the ENT at 4 months of age.
(F) Differences in cell-cell communication in the ENT at 8 months of age.
(G) Differences in cell-cell communication in the human RSC with MCI. For every panel in **Figure 4.5G-J**, nodes are colored by cell type, and edges are colored red to indicate an increase in signaling in either AD or MCI compared to controls, or blue to indicate a decrease in signaling in AD or MCI compared to controls.
(H) Differences in cell-cell communication in the human RSC with AD.
(I) Differences in cell-cell communication in the human ENT with MCI.
(J) Differences in cell-cell communication in the human ENT with AD.


A
RSC, 2 months
Oligodendrocytes
Astrocytes
GABAergic
Glutamatergic
Microglia
Endothelial
OPCs
B
RSC, 4 months
Oligodendrocytes
Astrocytes
GABAergic
Glutamatergic
Microglia
Endothelial
OPCs
C
RSC, 8 months
Oligodendrocytes
Astrocytes
GABAergic
Glutamatergic
Microglia
Endothelial
OPCs
D
ENT, 2 months
GABAergic
Astrocytes
OPCs
Glutamatergic
Microglia
Endothelial
Oligodendrocytes
E
ENT, 4 months
GABAergic
Astrocytes
OPCs
Glutamatergic
Microglia
Endothelial
Oligodendrocytes
F
ENT, 8 months
GABAergic
Astrocytes
OPCs
Glutamatergic
Microglia
Endothelial
Oligodendrocytes
G
RSC, MCI
Astrocytes
Oligodendrocytes
OPCs
Glutamatergic
Microglia
Endothelial
GABAergic
H
RSC, AD
Astrocytes
Oligodendrocytes
OPCs
Glutamatergic
Microglia
Endothelial
GABAergic
I
ENT, MCI
Oligodendrocytes
Astrocytes
Microglia
Glutamatergic
OPCs
Endothelial
GABAergic
J
ENT, AD
Oligodendrocytes
Astrocytes
Microglia
Glutamatergic
OPCs
Endothelial
GABAergic

differences (**Figure 4.5B, E**), like an average decrease in signaling to and from astrocytes in the RSC and an increase in astrocyte signaling in the ENT. But overall, the cellular communication changes inferred from CellChat at 4 months of age were more similar between the two brain regions than they appeared at 2 months. At 8 months of age, the two brain regions looked to be even more similar (**Figure 4.5C, F**).

When considering the results of CellChat analysis from human MCI and AD data (**Figure 4.5G-J**), we noted contrasts between the RSC in MCI (**Figure 4.5G**) and the ENT in MCI (**Figure 4.5I**). Oligodendrocytes appear to have a strong increase in signaling in the ENT, while microglia see a decrease. The opposite effects are seen in the RSC, along with other differences. Progression from MCI to AD in the RSC (**Figure 4.5G-H**) appears to result in an overall increase in signaling, while progression from MCI to AD in the ENT (**Figure 4.5I-J**) appears to result in a decrease.

To assess which signaling pathways were most influential in these communication differences across time and disease, we quantified the contribution of each signaling pathway detected within our datasets at significant levels (**Figure S4.5**). There was variability between pathways detected in mice (**Figure S4.5A**) and those detected in humans (**Figure S4.5B**). There was also an unfortunate sparsity in pathways detected in the RSC data from human tissue. When we considered only pathways detected in both species and in both brain regions, there was a lack of clear trends. For example, the EPHA signaling pathway met these requirements; in humans EPHA signaling was predicted to be downregulated in both MCI and AD in both brain regions. In mice, it was upregulated at 2 months, up in the ENT and down in the RSC at 4 months, and then appeared to switch to being up in the RSC and down in the ENT at 8 months. Because of inconsistent patterns of inferred communication, insights into specific pathways in the CellChat database were limited.

The same analysis was done in NeuronChat, but similar to the results from CellChat analysis, there was a lack of concordance between the mouse dataset and the human dataset (**Figure S4.6**), with many pathways (such as GABAergic signaling) being present in mice (**Figure S4.6A**) and not in humans, and a lack of glutamatergic signaling in two conditions where it was expected based on results from chapter 3 and differential gene expression analysis results from this very dataset. This lack of glutamatergic signaling in the RSC in MCI and in the ENT in AD (**Figure S4.6B**) was surprising to us, as we had previously seen the opposite in the ENT in chapter 3 of this dissertation (see chapter 3.3). This unexpected finding may be explained, unfortunately, by differences in QC metrics. The RSC samples from patients with MCI and the ENT samples from patients with AD in this dataset had lower average numbers of unique genes detected per cell and counts per cell (**Figure S4.1C-D**), which reflects the lower values of predicted glutamatergic signaling. As will be discussed in chapter 5, further sequencing of these brain regions should be done – not only more deeply, but also from a greater number of patients. While we were able to identify trends in overall differential gene expression (**Figure 4.3**), to get reliable information on specific pathways in human AD and MCI, more abundant and higher quality data is needed.

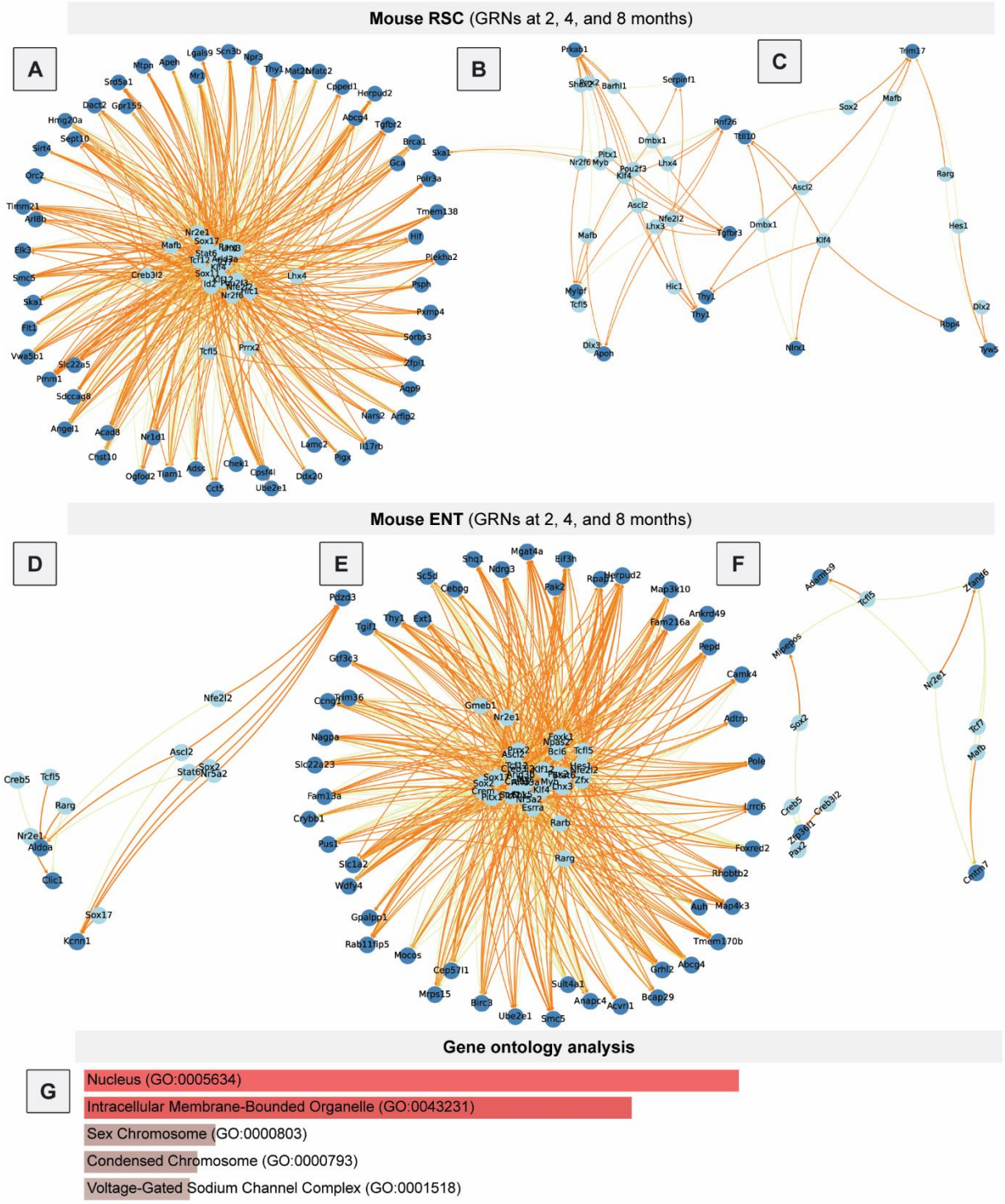
Comparison of gene regulatory networks inferred from multi-omic analysis

Following analysis of gene expression in both mouse and human data, we next turned to analysis of the snATAC-seq data we had available for mice. Due to findings highlighting differences in glutamatergic signaling, we focused this analysis on glutamatergic neurons. For each brain region and time point, we constructed a gene regulatory network (GRN) using DIRECT-NET[25] based only on genes that were upregulated in 5xFAD mice in glutamatergic neurons in that brain region at that time, resulting in 6 different GRNs: a 2-month-old RSC GRN, a 4-month-old RSC GRN, a 8-month-old RSC GRN, a 2-month-old ENT GRN, a 4-month-old ENT GRN, and a

8-month-old ENT GRN. In the construction of these networks, we further narrowed down what was included in the network by limiting potential transcription factor binding sites to peaks that were more accessible in 5xFAD mice compared to controls. The first thing we noted when observing the resulting GRNs, was that there were many more genes and transcription factors in the networks constructed based on 2-month-old RSC differences and 4-month-old ENT differences, even after increasing thresholds for inclusion in the network for just these two conditions (**Figure 4.6**). The RSC networks at 2-, 4- and 8-months of age (**Figure 4.6A-C**) became progressively smaller with each older time point. The ENT networks started small at 2 months (**Figure 4.6D**), with only 14 unique genes, ballooned up to hundreds at 4 months (only top 84 shown in **Figure 4.6E** to prevent too many overlaps), and then appeared to scale back down to 13 genes (**Figure 4.6F**). The size of the networks did not appear to be correlated to cell quantity or QC metrics, though we are not certain of the meaning behind the size differences. Based on these differences in networks, cellular activity data gathered earlier (**Figure 1E-H**), and differential gene expression analyses (**Figure 4.3-4.4**), we hypothesized that major changes seen in the RSC at 2 months of age could be being driven by epigenetic changes causing differential chromatin accessibility. There were not enough genes and transcription factors present in each network to do meaningful GO analysis for most time points, but fortunately the RSC at 2 months of age was an exception. While the top 4 terms most highly associated with the GRN constructed based on RSC glutamatergic neurons from 5xFAD mice at 4 months of age (**Figure 4.6A**) were not especially

Figure 4.6: 5xFAD-associated gene regulatory networks inferred from multiomic mouse data.
(A) GRN constructed based on upregulated genes and more highly accessible peaks in the multiomic data from glutamatergic neurons in the mouse RSC at 2 months of age.
(B) GRN constructed based on upregulated genes and more highly accessible peaks in the multiomic data from glutamatergic neurons in the mouse RSC at 4 months of age.
(C) GRN constructed based on upregulated genes and more highly accessible peaks in the multiomic data from glutamatergic neurons in the mouse RSC at 8 months of age.

(D) GRN constructed based on upregulated genes and more highly accessible peaks in the multiomic data from glutamatergic neurons in the mouse ENT at 2 months of age.
(E) GRN constructed based on upregulated genes and more highly accessible peaks in the multiomic data from glutamatergic neurons in the mouse ENT at 4 months of age.
(F) GRN constructed based on upregulated genes and more highly accessible peaks in the multiomic data from glutamatergic neurons in the mouse ENT at 8 months of age.
(G) GO analysis results from genes and transcription factors within the RSC 2-month GRN.



informative, the fifth term implicated voltage-gated sodium channel complexes (**Figure 4.6G**). Reflecting on the cellular activity data at this time (**Figure 4.1F**), gene expression differences in glutamatergic neurons at this time (**Figure 4.4A**), combined with these results, we propose that in 5xFAD mice, an increase in neuronal activity can potentially be explained by gene expression differences unique to the RSC at this time. The GRN constructed here implicates sodium channels as being regulated epigenetically, and it appears that this effect fades over time (due to the reduced size of GRNs at 4 and 8 months of age in **Figure 4.6B-C**).

4.4 Discussion

Here, we present a comparative analysis between disease-associated changes in the RSC and the ENT based on snRNA-seq data from individuals with AD and MCI, and from 5xFAD mice at varying time points. We first became interested in the RSC due to differences in RABV-tracing results implicating an increase in neuronal connectivity, and early hyperactivity in this brain region, and not the ENT, in 5xFAD mice (**Figure 4.1**). Following multi-omic sequencing in mice and either snRNA-seq in humans or collection of previously published data, we were able to delve into the molecular changes present in 5xFAD mice that may explain these neuronal activity changes, and assess similarity or differences between human disease and the mouse model within these two brain regions. After annotating clusters within the snRNA-seq data and determining that each of our datasets contained comparable cell types (**Figure 4.2**), we performed pairwise comparisons of disease (AD, MCI, or 5xFAD) to controls in each cell type and in each brain region. Linear regression was then applied to these changes in gene expression (**Figure 4.3**), to assess whether genes were disrupted by brain region in a similar fashion, or if region-specific changes were opposed. We found that at early time points in 5xFAD mice, region-specific changes were indeed opposed, as evidenced by negative correlations in every cell type (**Figure 4.3A**), but that

by 8 months of age correlations between brain regions had become positive. The differences between brain regions at 8 months of age in mice and in tissue from individuals with AD were similarly positively correlated (**Figure 4.3G, K**), suggesting that changes in gene expression later in disease are more similar between these two brain regions than at earlier disease stages. We next assessed what biological processes, molecular functions, or cellular components commonly dysregulated genes were most highly associated with (**Figure 4.4**), finding that synaptic proteins, glutamate receptors, and ion channels were significant terms common between the two groups of genes assessed: genes that were upregulated in AD in the RSC and the ENT (**Figure 4.4A-C**) and genes that were upregulated in MCI in the RSC but downregulated in MCI in the ENT (**Figure 4.4D-F**). We followed this focused analysis with a broader evaluation of cellular communication changes in AD, MCI, and 5xFAD (**Figure 4.5**). While some trends in communication were common across species, individual pathways had patterns of predicted signaling that were inconsistent between species. In the final component of this work which incorporated snATAC-seq data, we returned to a focused analysis of glutamatergic neurons. The GRNs we constructed based on upregulated genes and accessible sites within the genome gave insight into the potential role of epigenetic regulation in these cells, with the GRN at 2 months of age in the RSC being significantly associated with voltage-gated sodium channel complex genes.

Together, this work implicates early disruptions in glutamatergic neurons in the RSC in the development of AD from MCI, or in mice in progression from 2 to 4 months of age. We see changes in the RSC at 2 months that precede changes in the ENT in multiple contexts: cellular excitability, inferred cellular communication, and gene regulatory network analysis. A potential explanation for our observations would be that glutamatergic neurons in the RSC undergo changes relatively early in disease progression, and that these disruptions in activity in the RSC influence

the ENT via increased neuronal connections. Our results suggest that this early disruption could be epigenetically regulated, and that at least some aspects of these changes may be preserved in human AD. While many questions remain unanswered and further exploration is necessary, as will be discussed in chapter 5, the findings in this chapter are a promising step towards understanding the role of the RSC in AD, and how it may influence ENT disruptions.

4.5 Supplementary figures

Figure S4.1: QC metrics for snRNA-seq and snATAC-seq data after filtering.
(A) QC metrics from the mouse RSC snRNA-seq data. From left to right the distributions in red (5xFAD mice) and grey (control mice) for the following QC metrics are displayed in violin plots: nFeature RNA, or how many unique genes are detected in each cell, nCount RNA, or how many unique molecules were detected in each cell, and percent mt, or the percentage of detected counts that were mitochondrial genes. Distributions are split by age.
(B) QC metrics from the mouse ENT snRNA-seq data. The layout is the same as **S4.1A**.
(C) QC metrics from the patient donor RSC snRNA-seq data. From left to right the distributions in dark red (AD), red (MCI), and grey (Control, or cognitively normal) for the following QC metrics are displayed in violin plots: nFeature RNA, or how many unique genes are detected in each cell, nCount RNA, or how many unique molecules were detected in each cell, and percent mt, or the percentage of detected counts that were mitochondrial genes.
(D) QC metrics from the patient donor ENT snRNA-seq data. The layout is the same as **S4.1C**.
(E) QC metrics from the mouse RSC snATAC-seq data. From left to right the distributions in red (5xFAD mice) and grey (control mice) for the following QC metrics are displayed in violin plots: percent reads in peaks, counts in peaks, transcription start site enrichment, percent of reads being in a blacklist region, and nucleosome signal. Distributions are split by age.
(F) QC metrics from the mouse ENT snATAC-seq data. The layout is the same as **S4.1E**.

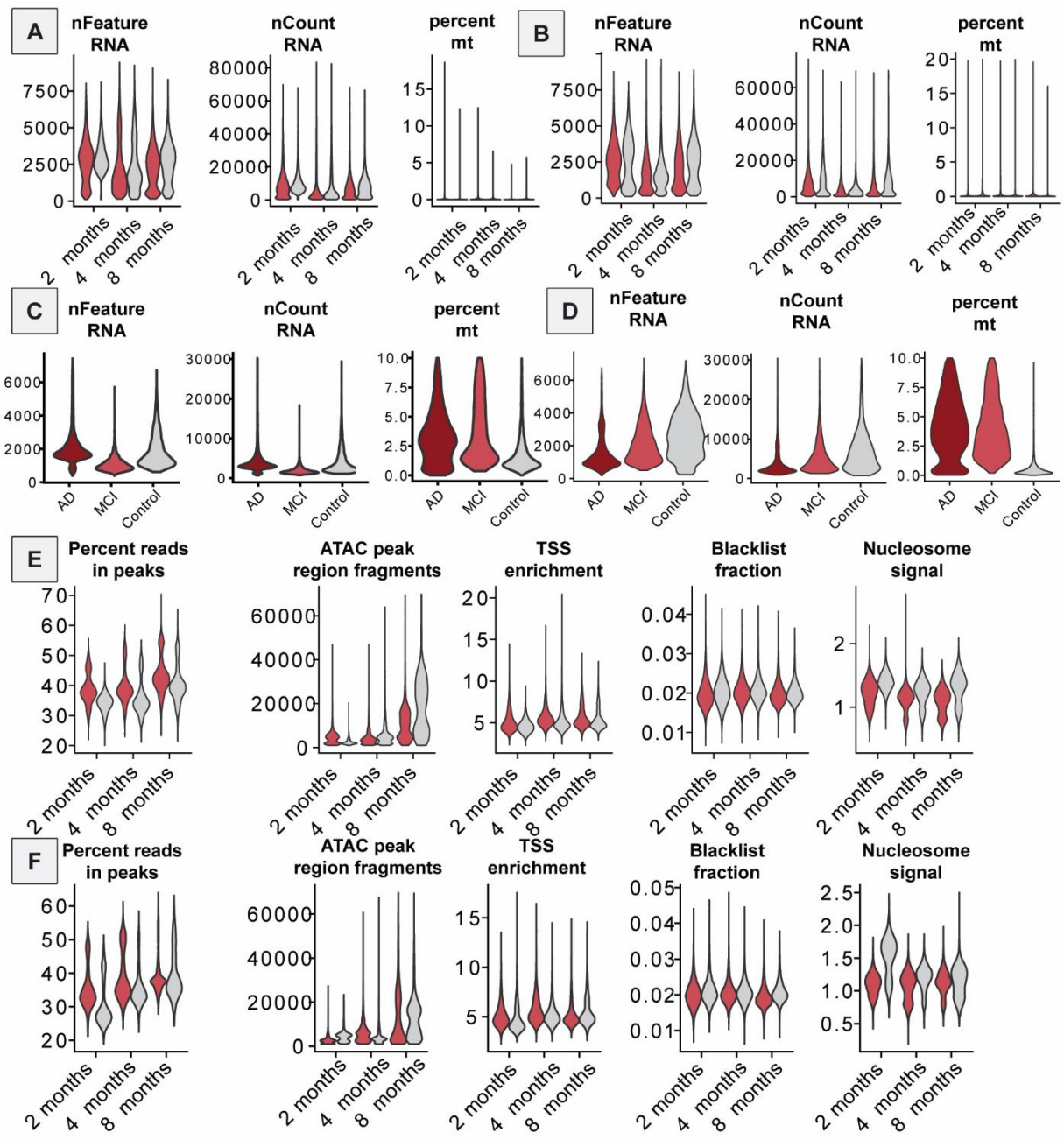

A
nFeature RNA
nCount RNA
percent mt
B
nFeature RNA
nCount RNA
percent mt
2 months
4 months
8 months
C
nFeature RNA
nCount RNA
percent mt
D
nFeature RNA
nCount RNA
percent mt
AD
MCI
Control
E
Percent reads in peaks
ATAC peak region fragments
TSS enrichment
Blacklist fraction
Nucleosome signal
F
Percent reads in peaks
ATAC peak region fragments
TSS enrichment
Blacklist fraction
Nucleosome signal

Figure S4.2: RSC cell type annotation.
(A) Dimensionality reduction of snRNA-seq data from the RSC in 5xFAD and control mice separated by age.
(B) Dimensionality reduction of snRNA-seq data from the RSC in 2-month-old, 4-month-old, and 8-month-old mice separated by genotype.
(C) Dimensionality reduction of snRNA-seq data from the RSC showing labels with 5 subclusters of GABAergic and 11 subclusters of glutamatergic neurons.
(D) Dimensionality reduction of snRNA-seq data from the RSC showing labels based on transferred annotations from the Allen Brain Atlas.
(E) Dotplot showing expression of both known marker genes, and genes that were markers of each neuronal subcluster in the annotated snRNA-seq data for the mouse RSC.
(F) Dimensionality reduction of snRNA-seq data from the human RSC separated by patient diagnosis.

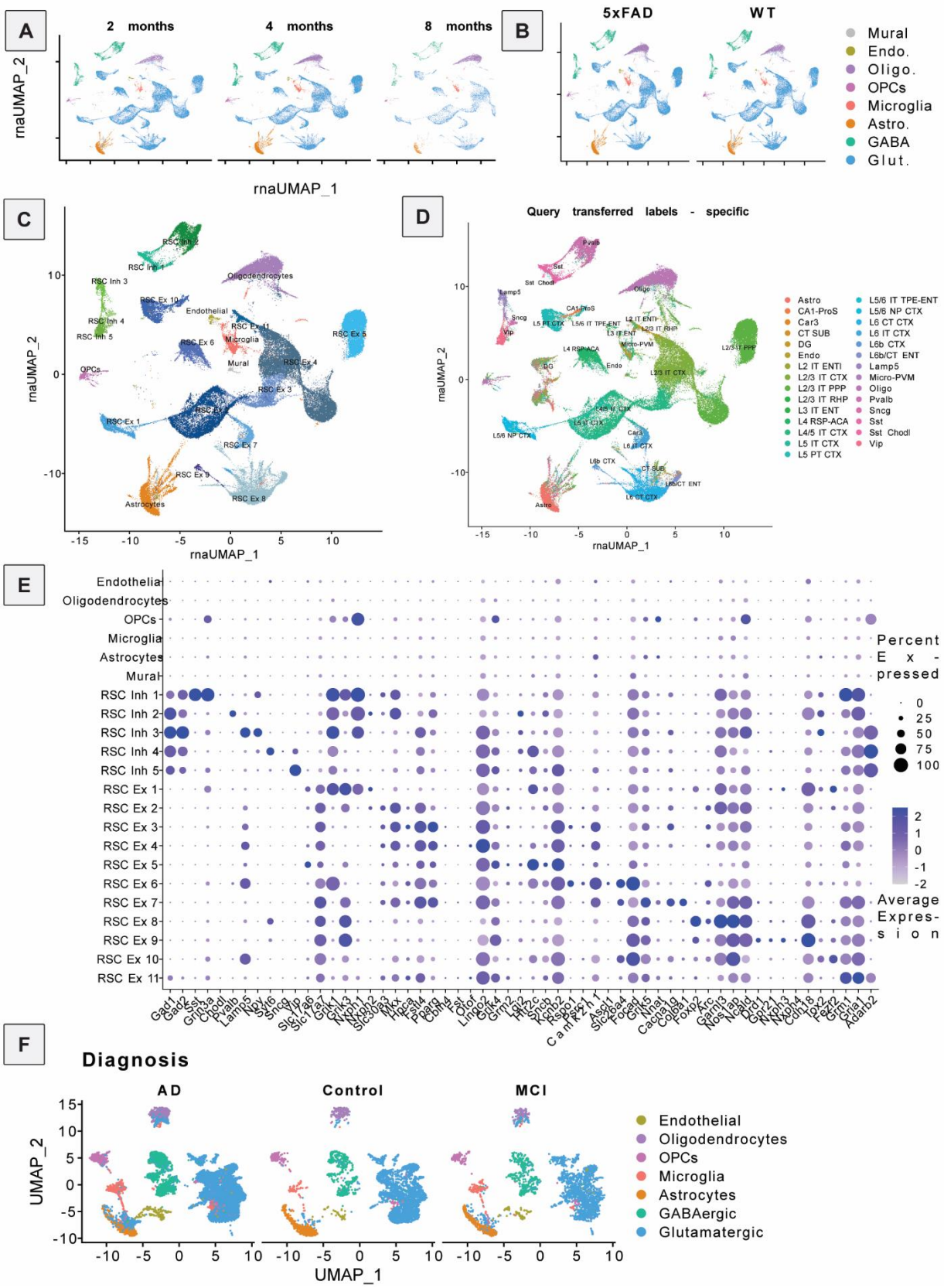

A
2 months
4 months
8 months
rnaUMAP_2
rnaUMAP_1
B
5xFAD
WT
Mural
Endo.
Oligo.
OPCs
Microglia
Astro.
GABA
Glut.
C
RSC Inh 1
RSC Inh 2
RSC Inh 3
RSC Inh 4
RSC Inh 5
RSC Ex 10
RSC Ex 11
RSC Ex 6
RSC Ex 5
RSC Ex 4
RSC Ex 3
RSC Ex 2
RSC Ex 1
RSC Ex 7
RSC Ex 9
RSC Ex 8
Oligodendrocytes
Endothelial
Microglia
Mural
OPCs
Astrocytes
D
Query transferred labels - specific
Astro
CA1-ProS
Car3
CT SUB
DG
Endo
L2 IT ENTI
L2/3 IT CTX
L2/3 IT PPP
L2/3 IT RHP
L3 IT ENT
L4 RSP-ACA
L4/5 IT CTX
L5 IT CTX
L5 PT CTX
L5/6 IT TPE-ENT
L5/6 NP CTX
L6 CT CTX
L6 IT CTX
L6b CTX
L6b/CT ENT
Lamp5
Micro-PVM
Oligo
Pvalb
Sncg
Sst
Sst Chodl
Vip
E
Endothelia
Oligodendrocytes
OPCs
Microglia
Astrocytes
Mural
RSC Inh 1
RSC Inh 2
RSC Inh 3
RSC Inh 4
RSC Inh 5
RSC Ex 1
RSC Ex 2
RSC Ex 3
RSC Ex 4
RSC Ex 5
RSC Ex 6
RSC Ex 7
RSC Ex 8
RSC Ex 9
RSC Ex 10
RSC Ex 11
Percent Expressed
Average Expression
F
Diagnosis
AD
Control
MCI
UMAP_2
UMAP_1
Endothelial
Oligodendrocytes
OPCs
Microglia
Astrocytes
GABAergic
Glutamatergic

Figure S4.3: ENT cell type annotation.
(A) Dimensionality reduction of snRNA-seq data from the ENT in 5xFAD and control mice separated by age.
(B) Dimensionality reduction of snRNA-seq data from the ENT in 2-month-old, 4-month-old, and 8-month-old mice separated by genotype.
(C) Dimensionality reduction of snRNA-seq data from the ENT showing labels with 5 subclusters of GABAergic and 11 subclusters of glutamatergic neurons.
(D) Dimensionality reduction of snRNA-seq data from the ENT showing labels based on transferred annotations from the Allen Brain Atlas.
(E) Dotplot showing expression of both known marker genes, and genes that were markers of each neuronal subcluster in the annotated snRNA-seq data for the mouse ENT.
(F) Dimensionality reduction of snRNA-seq data from the human ENT separated by patient diagnosis.

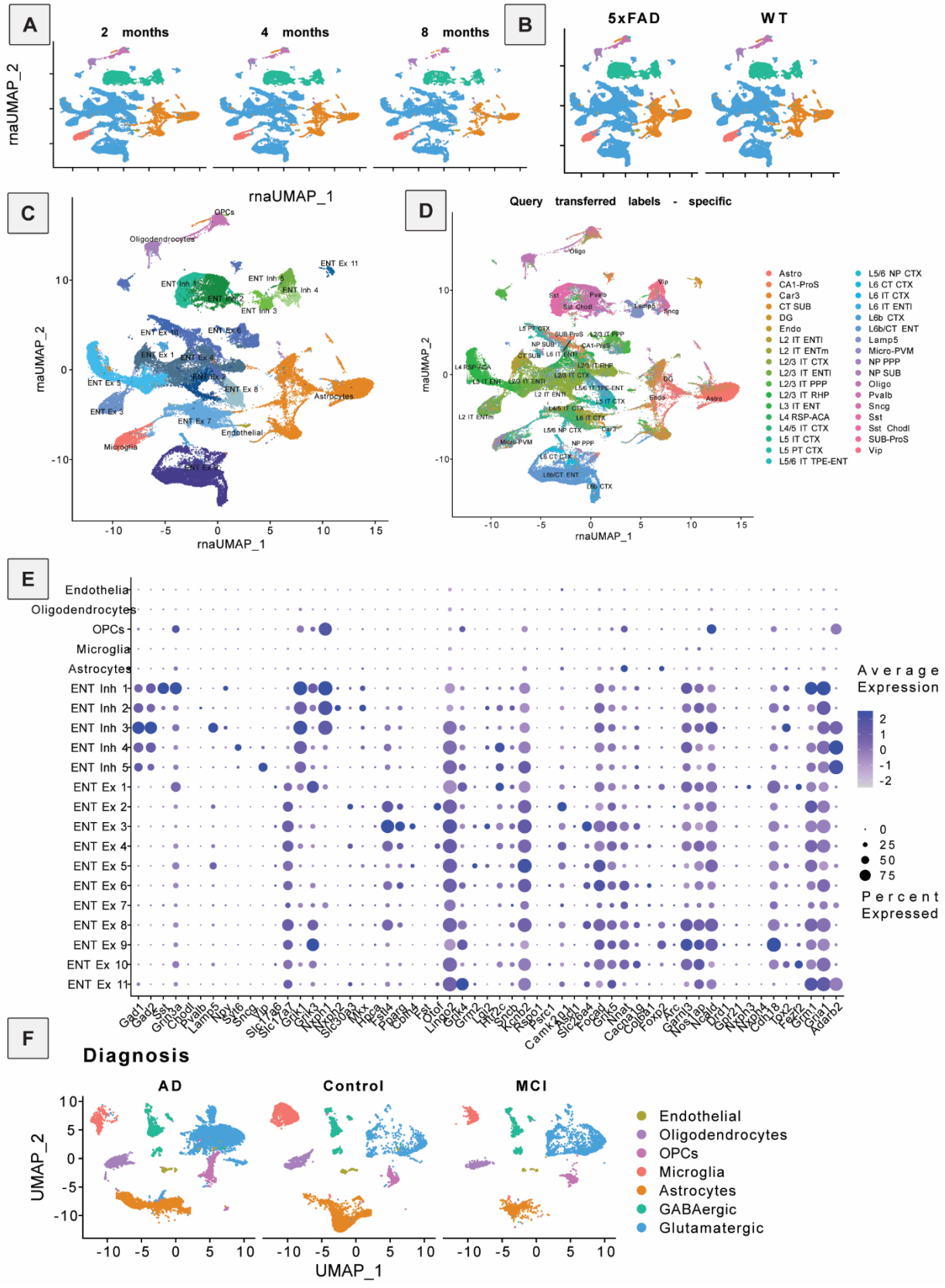

A
2 months
4 months
8 months
rnaUMAP_2
B
5xFAD
WT
C
rnaUMAP_1
rnaUMAP_2
OPCs
Oligodendrocytes
ENT Ex 11
ENT Inh 1
ENT Inh 5
ENT Inh 4
ENT Inh 2
ENT Inh 3
ENT Ex 10
ENT Ex 9
ENT Ex 1
ENT Ex 4
ENT Ex 2
ENT Ex 5
ENT Ex 8
Astrocytes
ENT Ex 3
ENT Ex 7
Endothelial
Microglia
ENT Ex 6
D
Query transferred labels - specific
Astro
CA1-ProS
Car3
CT SUB
DG
Endo
L2 IT ENTI
L2 IT ENTm
L2/3 IT CTX
L2/3 IT ENTI
L2/3 IT PPP
L2/3 IT RHP
L3 IT ENT
L4 RSP-ACA
L4/5 IT CTX
L5 IT CTX
L5 PT CTX
L5/6 IT TPE-ENT
L5/6 NP CTX
L6 CT CTX
L6 IT CTX
L6 IT ENTI
L6b CTX
L6b/CT ENT
Lamp5
Micro-PVM
NP PPP
NP SUB
Oligo
Pvalb
Sncg
Sst
Sst Chodl
SUB-ProS
Vip
E
Endothelia
Oligodendrocytes
OPCs
Microglia
Astrocytes
ENT Inh 1
ENT Inh 2
ENT Inh 3
ENT Inh 4
ENT Inh 5
ENT Ex 1
ENT Ex 2
ENT Ex 3
ENT Ex 4
ENT Ex 5
ENT Ex 6
ENT Ex 7
ENT Ex 8
ENT Ex 9
ENT Ex 10
ENT Ex 11
Gad1 Gad2 Sst Grin3a Chodl Pvalb Lamp5 Npy Syt6 Sncg Vip Slc17a6 Slc17a7 Grik1 Grik3 Nxph1 Nxph2 Slc30a3 Mkx Hpca Fstl4 Pappa Cdh4 Fst Otof Lingo2 Grik4 Grm2 Lgi2 Htr2c Sncb Kcnb2 Rspo1 Psrc1 Camk2a Syt1 Slc26a4 Foxp2 Gnb5 Nnat Cacna1g Col8a1 Foxp2 Arc Garnl3 Nos1ap Ncald Dpp10 Grin1 Nxph3 Nxph4 Cdh18 Tox2 Fezf2 Grin1 Gria1 Adarb2
Average Expression
2
1
0
-1
-2
0
25
50
75
Percent Expressed
F
Diagnosis
AD
Control
MCI
UMAP_2
UMAP_1
Endothelial
Oligodendrocytes
OPCs
Microglia
Astrocytes
GABAergic
Glutamatergic

Figure S4.4: DGE in glutamatergic neurons of the RSC and ENT in 5xFAD versus control mice.
(A) Volcano plot showing differential gene expression analysis results in 5xFAD versus control mice. These results are from glutamatergic neurons in the RSC at 2 months of age, and genes labeled are a subset of genes within the Postsynaptic Density (GO: 0014069) term.
(B) Volcano plot showing differential gene expression analysis results in 5xFAD versus control mice. These results are from glutamatergic neurons in the ENT at 2 months of age, and genes labeled are a subset of genes within the Postsynaptic Density (GO: 0014069) term.
(C) Volcano plot showing differential gene expression analysis results in 5xFAD versus control mice. These results are from glutamatergic neurons in the RSC at 8 months of age, and genes labeled are a subset of genes within the Postsynaptic Density (GO: 0014069) term.
(D) Volcano plot showing differential gene expression analysis results in 5xFAD versus control mice. These results are from glutamatergic neurons in the ENT at 8 months of age, and genes labeled are a subset of genes within the Postsynaptic Density (GO: 0014069) term.

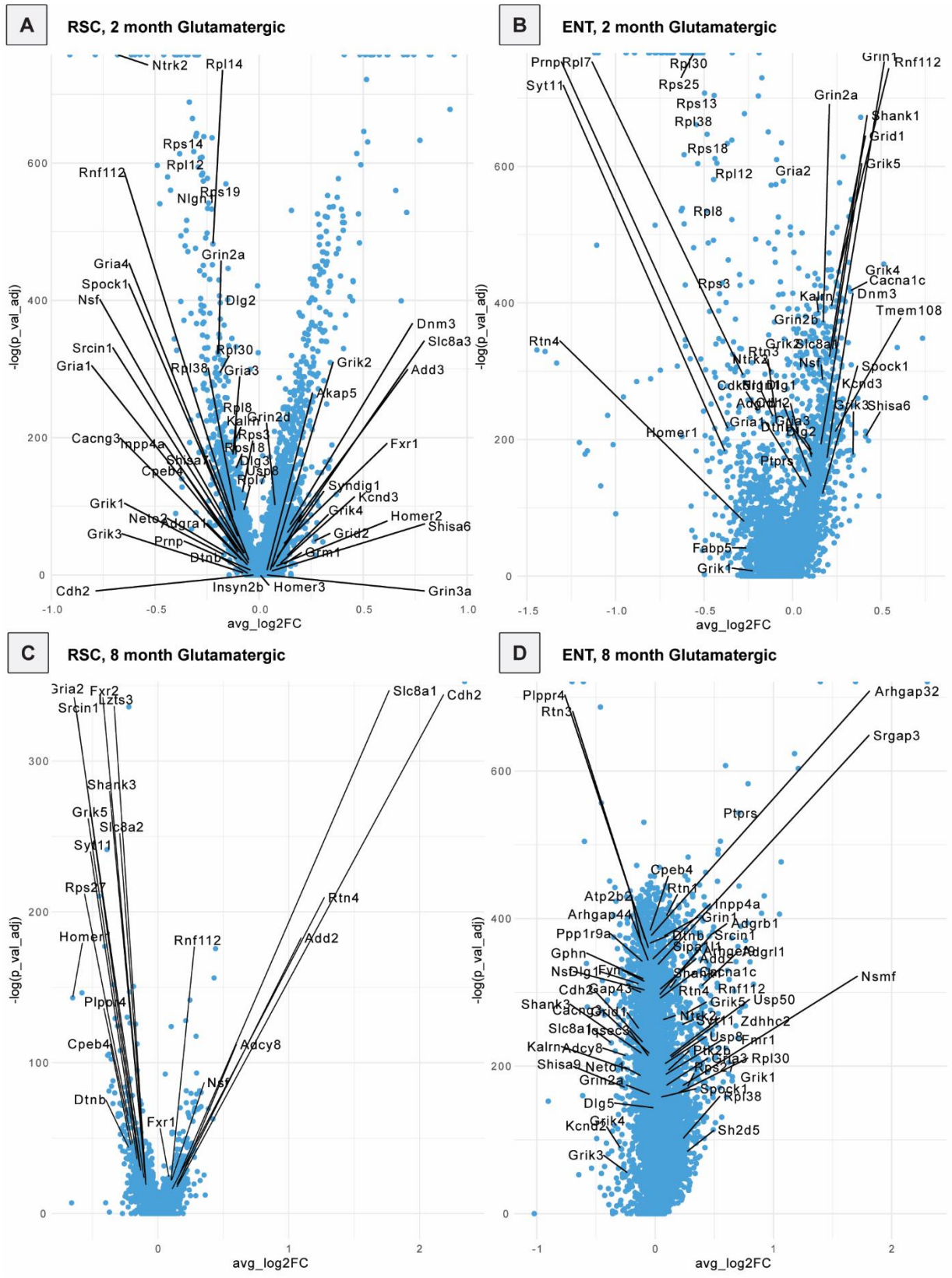
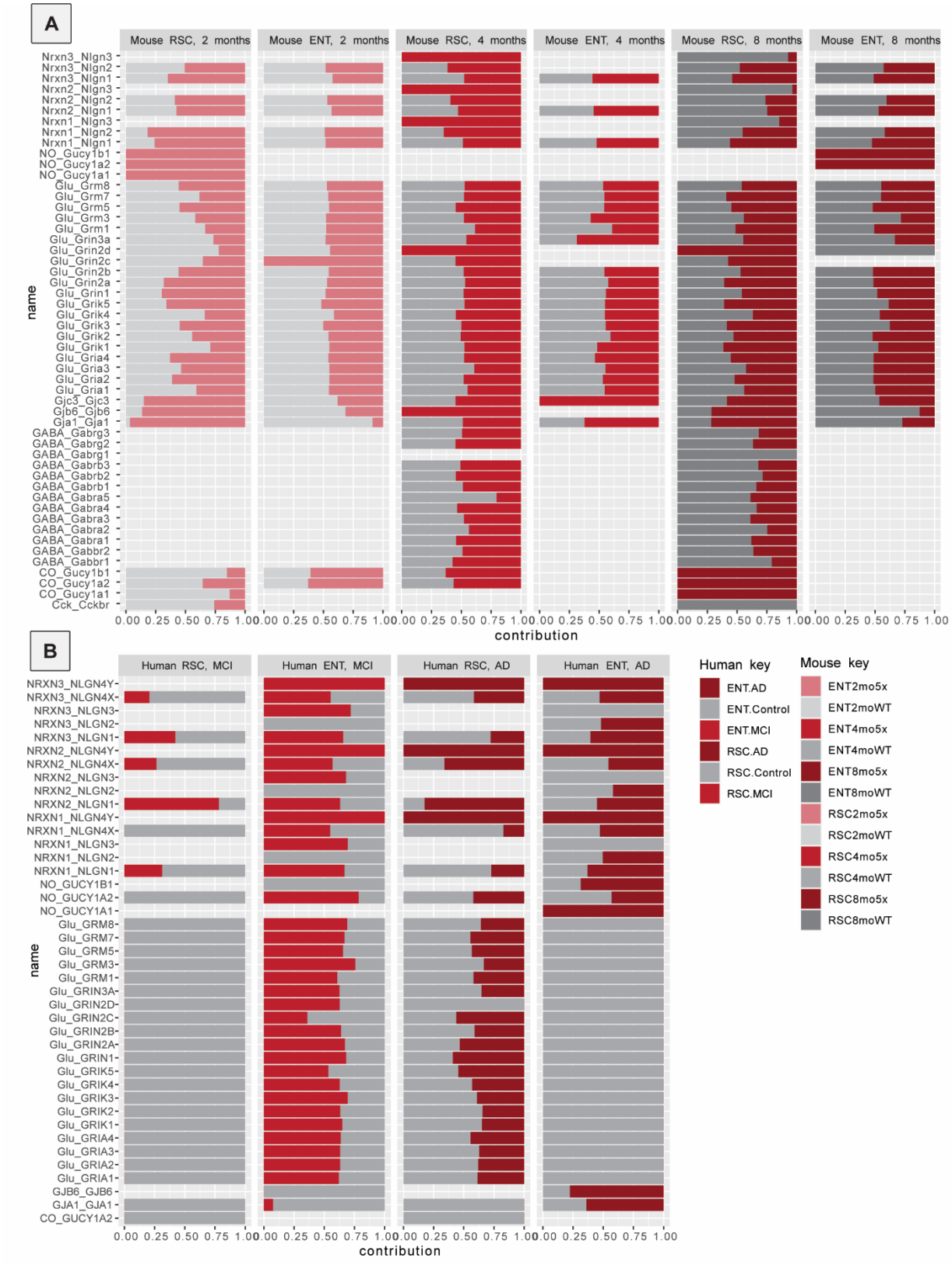

A
RSC, 2 month Glutamatergic
B
ENT, 2 month Glutamatergic
C
RSC, 8 month Glutamatergic
D
ENT, 8 month Glutamatergic
-log(p_val_adj)
avg_log2FC

Figure S4.5: CellChat signaling pathway differences in 5xFAD mice and patients.
(A) Bar plots showing differences in predicted signaling for each pathway in the CellChat database identified within the mouse data. Pairwise comparisons of contribution to signaling between 5xFAD and control (WT) mice are displayed for each time point in each brain region.
(B) Bar plots showing differences in predicted signaling for each pathway in the CellChat database identified within the human data. Pairwise comparisons of contribution to signaling between either AD and control patients or MCI and control patients are displayed for each brain region.


A
Mouse RSC, 2 months
Mouse ENT, 2 months
Mouse RSC, 4 months
Mouse ENT, 4 months
Mouse RSC, 8 months
Mouse ENT, 8 months
name
contribution
B
Human RSC, MCI
Human ENT, MCI
Human RSC, AD
Human ENT, AD
Human key
ENT.AD
ENT.Control
ENT.MCI
RSC.AD
RSC.Control
RSC.MCI
Mouse key
ENT2mo5x
ENT2moWT
ENT4mo5x
ENT4moWT
ENT8mo5x
ENT8moWT
RSC2mo5x
RSC2moWT
RSC4mo5x
RSC4moWT
RSC8mo5x
RSC8moWT

Figure S4.6: NeuronChat signaling pathway differences in 5xFAD mice and patients.
(A) Bar plots showing differences in predicted signaling for each pathway in the NeuronChat database identified within the mouse data. Pairwise comparisons of contribution to signaling between 5xFAD and control (WT) mice are displayed for each time point in each brain region.
(B) Bar plots showing differences in predicted signaling for each pathway in the NeuronChat database identified within the human data. Pairwise comparisons of contribution to signaling between either AD and control patients or MCI and control patients are displayed for each brain region.


A
Mouse RSC, 2 months
Mouse ENT, 2 months
Mouse RSC, 4 months
Mouse ENT, 4 months
Mouse RSC, 8 months
Mouse ENT, 8 months
name
contribution
B
Human RSC, MCI
Human ENT, MCI
Human RSC, AD
Human ENT, AD
Human key
ENT.AD
ENT.Control
ENT.MCI
RSC.AD
RSC.Control
RSC.MCI
Mouse key
ENT2mo5x
ENT2moWT
ENT4mo5x
ENT4moWT
ENT8mo5x
ENT8moWT
RSC2mo5x
RSC2moWT
RSC4mo5x
RSC4moWT
RSC8mo5x
RSC8moWT

4.6 References

1. Vann, S. D., Aggleton, J. P. & Maguire, E. A. What does the retrosplenial cortex do? *Nature Reviews Neuroscience* vol. 10 792–802 Preprint at https://doi.org/10.1038/nrn2733 (2009).
2. Trask, S. & Fournier, D. I. Examining a role for the retrosplenial cortex in age-related memory impairment. *Neurobiol Learn Mem* **189**, (2022).
3. Walsh, C. *et al.* β Bursting in the Retrosplenial Cortex Is a Neurophysiological Correlate of Environmental Novelty Which Is Disrupted in a Mouse Model of Alzheimer's Disease. *Journal of Neuroscience* **42**, 7094–7109 (2022).
4. Ziontz, J., Adams, J. N., Harrison, T. M., Baker, S. L. & Jagust, W. J. Hippocampal connectivity with retrosplenial cortex is linked to neocortical tau accumulation and memory function. *Journal of Neuroscience* **41**, 8839 (2021).
5. Wickersham, I. R. *et al.* Monosynaptic Restriction of Transsynaptic Tracing from Single, Genetically Targeted Neurons. *Neuron* **53**, 639–647 (2007).
6. Oakley, H. *et al.* Intraneuronal β-amyloid aggregates, neurodegeneration, and neuron loss in transgenic mice with five familial Alzheimer's disease mutations: Potential factors in amyloid plaque formation. *Journal of Neuroscience* **26**, 10129–10140 (2006).
7. Ye, Q. *et al.* Hippocampal neural circuit connectivity alterations in an Alzheimer's disease mouse model revealed by monosynaptic rabies virus tracing. *Neurobiol Dis* **172**, (2022).
8. Ye, Q. *et al.* Monosynaptic Rabies Tracing Reveals Sex- and Age-Dependent Dorsal Subiculum Connectivity Alterations in an Alzheimer's Disease Mouse Model. *Journal of Neuroscience* **44**, (2024).
9. Da Silva, S. V. *et al.* Early synaptic deficits in the APP/PS1 mouse model of Alzheimer's disease involve neuronal adenosine A2A receptors. *Nat Commun* **7**, (2016).
10. Schmid, L. C. *et al.* Dysfunction of Somatostatin-Positive Interneurons Associated with Memory Deficits in an Alzheimer's Disease Model. *Neuron* **92**, 114–125 (2016).
11. Renier, N. *et al.* iDISCO: A Simple, Rapid Method to Immunolabel Large Tissue Samples for Volume Imaging. *Cell* **159**, 896–910 (2014).
12. Beier, K. T. *et al.* Rabies screen reveals GPe control of cocaine-triggered plasticity. *Nature* **549**, 345–350 (2017).
13. Lerner, T. N. *et al.* Intact-Brain Analyses Reveal Distinct Information Carried by SNc Dopamine Subcircuits. *Cell* **162**, 635–647 (2015).
14. Gunaydin, L. A. *et al.* Natural Neural Projection Dynamics Underlying Social Behavior. *Cell* **157**, 1535–1551 (2014).
15. Smith, A. M. *et al.* Diverse human astrocyte and microglial transcriptional responses to Alzheimer's pathology. *Acta Neuropathol* **143**, 75–91 (2022).
16. Grubman, A. *et al.* A single-cell atlas of entorhinal cortex from individuals with Alzheimer's disease reveals cell-type-specific gene expression regulation. *Nat Neurosci* **22**, 2087–2097 (2019).
17. Hodge, R. D. *et al.* Conserved cell types with divergent features in human versus mouse cortex. *Nature* **573**, 61–68 (2019).
18. Yao, Z. *et al.* A taxonomy of transcriptomic cell types across the isocortex and hippocampal formation. *Cell* **184**, 3222-3241.e26 (2021).
19. Finak, G. *et al.* MAST: A flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol* **16**, (2015).
20. Stuart, T., Srivastava, A., Madad, S., Lareau, C. A. & Satija, R. Single-cell chromatin state analysis with Signac. *Nat Methods* **18**, 1333–1341 (2021).

21. Chen, E. Y. *et al. Enrichr: Interactive and Collaborative HTML5 Gene List Enrichment Analysis Tool*. http://amp.pharm.mssm.edu/Enrichr. (2013).
22. Kuleshov, M. V. *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* **44**, W90–W97 (2016).
23. Jin, S. *et al.* Inference and analysis of cell-cell communication using CellChat. *Nat Commun* **12**, (2021).
24. Zhao, W., Johnston, K. G., Ren, H., Xu, X. & Nie, Q. Inferring neuron-neuron communications from single-cell transcriptomics through NeuronChat. *Nat Commun* **14**, (2023).
25. Zhang, L., Zhang, J. & Nie, Q. *DIRECT-NET: An Efficient Method to Discover Cis-Regulatory Elements and Construct Regulatory Networks from Single-Cell Multiomics Data*. *Sci. Adv* vol. 8 https://www.science.org (2022).
26. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, (2008).
27. Korsunsky, I., Nathan, A., Millard, N. & Raychaudhuri, S. Presto scales Wilcoxon and auROC analyses to millions of observations. *bioRxiv* (2019) doi:10.1101/653253.
28. Ntranos, V., Yi, L., Melsted, P. & Pachter, L. A discriminative learning approach to differential expression analysis for single-cell RNA-seq. *Nat Methods* **16**, 163–166 (2019).
29. Hagberg hagberg, A. A., -Los, lanlgov, Schult, D. A. & Swart swart, P. J. *Exploring Network Structure, Dynamics, and Function Using NetworkX*. http://conference.scipy.org/proceedings/SciPy2008/paper_2 (2008).
30. Park, H. *et al.* Single-cell RNA-sequencing identifies disease-associated oligodendrocytes in male APP NL-G-F and 5XFAD mice. *Nat Commun* **14**, 802 (2023).
31. Safaiyan, S. *et al.* White matter aging drives microglial diversity. *Neuron* **109**, 1100-1117.e10 (2021).
32. Habib, N. *et al.* Disease-associated astrocytes in Alzheimer's disease and aging. *Nat Neurosci* **23**, 701–706 (2020).
33. Kiani Shabestari, S. *et al.* Absence of microglia promotes diverse pathologies and early lethality in Alzheimer's disease mice. *Cell Rep* **39**, (2022).
34. Fatmi, M. K. *et al.* Single-Cell RNA-seq reveals transcriptomic modulation of Alzheimer's disease by activated protein C. *Aging* **16**, 3137–3159 (2024).
35. Keren-Shaul, H. *et al.* A Unique Microglia Type Associated with Restricting Development of Alzheimer's Disease. *Cell* **169**, 1276-1290.e17 (2017).

CHAPTER 5: Conclusion and future directions

5.1 Discussion

The work presented in this dissertation combines results from multi-omic analysis with findings from neuronal tracing and other experimental methods to explore cell-type and brain-region specific changes in AD over time. Early changes in glutamatergic neurons in the RSC compared to the ENT are explored, and commonalities between changes in the mouse model and the patient data are identified.

Chapter 1 presented an overview of background information and previous research performed in AD and in mouse models of AD. The history and impact of mouse models in AD research was summarized, as well as the growing insights yielded from transcriptomic analysis. The mouse model we chose to use, 5xFAD, was described in detail, and neuronal tracing methods were reviewed. Previous usage of retrograde RABV tracing in mouse models of AD were also described, though these studies focused on the hippocampus rather than the ENT.

In chapter 2, the utility of RABV-tracing was explored in a different context: how inputs change after exposure to addictive drugs. This work showed that RABV-tracing results reflect differences in drug administration and target cell types, and that basal calcium ion channel expression in a brain region correlates to RABV-infected cells in that brain region following drug exposure. These findings combined suggest that RABV tracing differences can be indicative of cellular excitability changes in input brain regions.

In chapter 3 we characterized cellular communication changes in the ENT that were common to both 5xFAD mice and AD patients. At early ages in the mouse model, commonalities that persisted were sparse, but at the later time point of 8 months of age, major changes in

glutamatergic signaling were identified that aligned with changes in the human ENT, and were reflected in a GRN constructed based on multi-omic data.

Chapter 4 built upon the findings from chapters 2 and 3, first exploring RABV tracing in 5xFAD mice compared to controls to determine inputs to ENT neurons affected in early AD pathology. The resulting analysis identified the RSC as an important input region to the ENT, and further experimentation showed hyperexcitability in the RSC as early as 2 months of age, while this was only seen in the ENT at 8 months of age in 5xFAD mice. The subsequent sequencing experiments and analyses performed to characterize transcriptomic changes in these mouse brain regions suggested that differences early in disease progressively become more similar in the RSC and ENT as pathology worsens, and this effect is observed in the human RSC and ENT as well.

While there were limitations to this work, some of which are outlined below, the findings cumulatively show the importance of brain-region specific analysis in the study of AD progression, and will hopefully result in further interest in and exploration of the RSC and hyperexcitability in AD.

5.2 Future directions

Increase in number of subjects for studies in both mouse models and in patients with AD, MCI

In chapters 3 and 4 we utilized data collected from both mouse tissue and patient donor tissue. There were a multitude of conditions: 5xFAD mice at 2 months of age, 5xFAD mice at 4 months of age, 5xFAD mice at 8 months of age, C57BL/6 mice at 2 months of age, C57BL/6 mice at 4 months of age, C57BL/6 mice at 8 months of age, patients with AD, patients with MCI, and patients without cognitive impairment. Each of these conditions was studied in two distinct brain regions, the ENT and the RSC, essentially doubling the number of conditions for which replicates

were necessary. While the number of cells and depth of sequencing data were satisfactory for our purposes, there was some sparsity given limited time and resources, especially in 5xFAD mice at 8 months of age, and in samples of the human RSC (data from samples of the human ENT was supplemented with previously published data). For these conditions in particular, further collection of high-quality sequencing data would ensure that our findings are robust to differences between individuals and allow for further investigation into genes that are lowly expressed but still significant, and deeper probing into changes in rare cell types that were previously under-represented.

Sex differences in AD

In a similar vein as the previous discussion point, increasing the number of subjects utilized would also allow for exploration of sex differences in AD. All of the sequencing data from 5xFAD mice in chapters 3 and 4 of this work was from female mice. The data from female 5xFAD mice may not be generalizable to all 5xFAD mice, especially considering the previously characterized sex differences in this model[1]. Analyses done in chapters 3 and 4 ought to be repeated with male 5xFAD mice, and results compared and contrasted. Postmortem tissue from the ENT of patients analyzed in chapter 3 was also exclusively from female patients. Chapter 4 included data from both male and female patients, but sex differences were not the focus of this study and were not considered, partially due to a small number of samples. Future work should include data from a greater number of patients, with characterization of sex differences in both the mouse model and in patient data being performed.

Mouse models of tauopathy

In addition to modeling characteristics of AD associated with amyloid pathology, it is important to consider the effects of other factors, namely tau. Tau pathology was an inherent part

of the data collected from patient donors, but Braak staging information, while available, was not a part of our analyses. Chapters 3 and 4 of this dissertation utilized data from 5xFAD mice, which do not model tau pathology. While we found some similarities between the mouse model and patient data, there were differences in gene expression and predicted cellular communication patterns that did not align across species. The gene expression differences in mouse models that reflect tau pathology, or both tau and amyloid pathology, need further investigation. For this reason, future research should include snRNA-seq data from mouse models of tauopathy such as the P301S transgenic model, the PS19 mouse[2] or the rTg4510 mouse[3]. Further, models that combine both tau and amyloid, like the TauPS2APP triple transgenic mouse model[4] for example, should be considered in future work, as both tau and amyloid play a role in AD.

Further investigation via proteomics and metabolomics

The work in this dissertation is largely based on findings from snRNA-seq data. While this information is of value, it does not always reflect the translation of detected mRNAs to proteins, with gene expression as assessed by RNA-seq previously being shown to have only a moderate correlation to protein levels[5]. With protein quantification methods improving in scale due to advances in next-generation sequencing and mass spectrometry[6], and new methods such as CITE-seq[7] allowing for integration of transcriptomic and protein measurements, verification of findings from previous large scale transcriptomic studies is now more feasible and would be a valuable inclusion in further study. Additionally, beyond proteins, other macromolecules ought to be considered in the study of AD; lipids are of particular interest, partly due to the influence of the *APOE* gene in AD risk. The *APOE4* allele is a strong genetic risk factor in the development of AD, and APOE is known to regulate lipid metabolism[8]. Further, the plasma lipidome has previously been shown to change in AD patients compared to healthy controls[9]. Assessment of the

proteome and lipidome in AD could yield insights into disease-associated changes that are not reflected in transcriptomic data, and would be valuable additions to further study in validation efforts or even exploratory analysis.

Consideration of additional brain regions

To gain a comprehensive understanding of AD, neural circuits and the corresponding changes in gene regulation, we must look beyond the RSC and ENT. While the work done in this dissertation on the ENT and the RSC yielded insights into the cellular excitability changes in these regions throughout the stages of AD, further analysis and comparison to brain regions such as the hippocampus and relevant inputs could be just as fruitful. The previously published RABV tracing work discussed in chapter 1 focused on inputs to various regions and neuron types within the hippocampus[10–13], showing differences in a variety of mouse models worth further exploration. From our own work, the neuronal tracing results shown in chapter 4 appear to show disruptions in this area as well, as the subiculum, a subregion within the hippocampus, was the only input to the ENT to have significantly increased in 5xFAD mice other than the RSC. Exploration of cellular activity, along with molecular changes as assessed by transcriptomic and epigenomic data from hippocampal and other brain regions, would be a natural next step based on these results.

Recent work has delved into the differences by brain region apparent in AD, with continued efforts underway to construct a multi-region atlas of AD including single-cell data from the ENT, hippocampus, anterior thalamus, angular gyrus, midtemporal cortex, and prefrontal cortex[14]. There is ongoing work in the Beier lab to more deeply sequence the RSC, ENT and sequence the hippocampus, and with the increased availability and improvement of sequencing technologies, a brain-wide atlas of individuals with AD or MCI compared to non-cognitively impaired individuals is becoming a more plausible endeavor as time goes on. Given our findings in the RSC, a deeper

understanding of how AD affects other brain regions that have historically been overlooked, particularly at early stages of disease, would be invaluable in the study of AD progression.

5.3 References

1. Oblak, A. L. *et al*. Comprehensive Evaluation of the 5XFAD Mouse Model for Preclinical Testing Applications: A MODEL-AD Study. *Front Aging Neurosci* **13**, (2021).
2. Yoshiyama, Y. *et al*. Synapse Loss and Microglial Activation Precede Tangles in a P301S Tauopathy Mouse Model. *Neuron* **53**, 337–351 (2007).
3. Ramsden, M. *et al*. Age-Dependent Neurofibrillary Tangle Formation, Neuron Loss, and Memory Impairment in a Mouse Model of Human Tauopathy (P301L). *The Journal of Neuroscience* **25**, 10637–10647 (2005).
4. Grueninger, F. *et al*. Phosphorylation of Tau at S422 is enhanced by Aβ in TauPS2APP triple transgenic mice. *Neurobiol Dis* **37**, 294–306 (2010).
5. Fu, X. *et al*. Estimating accuracy of RNA-Seq and microarrays with proteomics. *BMC Genomics* **10**, 161 (2009).
6. Bennett, H. M., Stephenson, W., Rose, C. M. & Darmanis, S. Single-cell proteomics enabled by next-generation sequencing or mass spectrometry. *Nat Methods* **20**, 363–374 (2023).
7. Stoeckius, M. *et al*. Simultaneous epitope and transcriptome measurement in single cells. *Nat Methods* **14**, 865–868 (2017).
8. Yang, L. G., March, Z. M., Stephenson, R. A. & Narayan, P. S. Apolipoprotein E in lipid metabolism and neurodegenerative disease. *Trends in Endocrinology & Metabolism* **34**, 430–445 (2023).
9. Liu, Y. *et al*. Plasma lipidome is dysregulated in Alzheimer's disease and is associated with disease risk genes. *Transl Psychiatry* **11**, (2021).
10. Ye, Q. *et al*. Hippocampal neural circuit connectivity alterations in an Alzheimer's disease mouse model revealed by monosynaptic rabies virus tracing. *Neurobiol Dis* **172**, (2022).
11. Ye, Q. *et al*. Monosynaptic Rabies Tracing Reveals Sex- and Age-Dependent Dorsal Subiculum Connectivity Alterations in an Alzheimer's Disease Mouse Model. *Journal of Neuroscience* **44**, (2024).
12. Da Silva, S. V. *et al*. Early synaptic deficits in the APP/PS1 mouse model of Alzheimer's disease involve neuronal adenosine A2A receptors. *Nat Commun* **7**, (2016).
13. Schmid, L. C. *et al*. Dysfunction of Somatostatin-Positive Interneurons Associated with Memory Deficits in an Alzheimer's Disease Model. *Neuron* **92**, 114–125 (2016).
14. Mathys, H. *et al*. Single-cell multiregion dissection of Alzheimer's disease. *Nature* (2024) doi:10.1038/s41586-024-07606-7.