

---

Dynamical systems analysis to  
characterize transcriptomic and  
neuroimaging changes in Alzheimer's  
disease and Parkinson's disease

---

Quadri Adewale  
Integrated Program in Neuroscience  
McGill University, Montreal

August 2024

*A thesis submitted to McGill University in partial fulfilment  
of the requirements of the degree of Doctor of Philosophy*

©Quadri Adewale, 2024

## Abstract

Alzheimer's disease (AD) and Parkinson's disease (PD) are the two most common neurodegenerative disorders, and they are characterized by clinical heterogeneity and multiple pathological processes. There is therefore a critical need to understand how the interactions between the various disease-related processes engender symptom manifestation and heterogeneity. Recent scientific advances have facilitated the measurement of disease-associated biological processes using RNA-sequencing and neuroimaging methods. Leveraging these advances, this thesis aimed to study the multiscale and multifactorial processes underlying AD and PD. Chapter 1 provides the motivation for this thesis and reviews relevant literature on the pathological processes underlying AD and PD. Various disease progression modelling methods used in studying those pathological processes are also discussed. Chapter 2 develops a novel data-driven dynamical systems model to study multiscale brain changes in aging and AD, which were observed to share some underlying biological mechanisms, with AD having more dysregulated processes. Chapter 3 further applies the developed model to study PD pathological processes and their relation to clinical presentations and physical activity. We found that different biological mechanisms underlie heterogeneity in PD symptom manifestation. Furthermore, the clinical utility of the model is demonstrated via *in silico* perturbation to reveal putative PD drugs. Chapter 4 investigates AD pathological changes across 6 different brain cell types by leveraging dynamical systems model of single-cell transcriptomics. Accelerated cell changes were observed in AD compared to normal aging. Chapter 5 discusses the contributions of this thesis, the limitations of its approach, and the suggestions for future work. Together, this work constitutes a disease-agnostic and multiscale data-driven approach that provides comprehensive insights into the complex multifactorial pathogenesis of AD and PD, unravels key biological modulators of

physical activity and clinical deterioration, and serves as a computational tool for personalized drug discovery.

## Résumé

La maladie d'Alzheimer (MA) et la maladie de Parkinson (MP) sont les deux troubles neurodégénératifs les plus courants et se caractérisent par une hétérogénéité clinique et de multiples processus pathologiques. Il existe donc un besoin crucial de comprendre comment les interactions entre les différents processus liés aux maladies engendrent la manifestation et l'hétérogénéité des symptômes. Les progrès scientifiques récents ont facilité la mesure des processus biologiques associés aux maladies à l'aide de méthodes de séquençage d'ARN et de neuroimagerie. Tirant parti de ces avancées, cette thèse visait à étudier les processus multi-échelles et multifactoriels sous-jacents à la MA et à la PD. Le chapitre 1 fournit la motivation de cette thèse et passe en revue la littérature pertinente sur les processus pathologiques sous-jacents à la MA et à la MP. Diverses méthodes de modélisation de la progression de la maladie utilisées dans pour étudier ces processus pathologiques sont également discutées. Le chapitre 2 développe un nouveau modèle de systèmes dynamiques basé sur des données pour étudier les changements cérébraux multi-échelles liés au vieillissement et à la MA, qui partagent certains mécanismes biologiques sous-jacents, la MA ayant des processus plus dérégulés. Le chapitre 3 applique en outre le modèle développé pour étudier les processus pathologiques de la MP et leur relation avec les présentations cliniques et l'activité physique. Nous avons constaté que différents mécanismes biologiques sont à l'origine de l'hétérogénéité de la manifestation des symptômes de la MP. De plus, l'utilité clinique du modèle est démontrée via une perturbation *in silico* pour révéler des médicaments putatifs contre la MP. Le chapitre 4 étudie les changements pathologiques de la MA dans 6 types de cellules cérébrales différents en tirant parti du modèle de systèmes dynamiques de transcriptomique unicellulaire. Des changements cellulaires accélérés ont été observés dans la MA par rapport au vieillissement normal. Le chapitre 5 discute des contributions de cette thèse, des limites de son approche et des suggestions de travaux futurs. Ensemble, ces travaux constituent une approche

indépendante de la maladie et basée sur des données multi-échelles qui fournit des informations complètes sur la pathogenèse multifactorielle complexe de la MA et de la MP, dévoile les principaux modulateurs biologiques de l'activité physique et de la détérioration clinique et sert d'outil informatique pour la découverte de médicaments personnalisés.

# Contents

Abstract.....	i
Résumé.....	iii
Contents .....	v
List of Figures .....	x
List of Tables .....	xii
List of Abbreviations .....	xiii
Acknowledgements.....	xiv
Contribution of authors .....	xvi
Contributions to knowledge.....	xviii
Chapter 1 : Introduction and Literature Review .....	1
1.1 Rationale and Objectives .....	1
1.1.1 Rationale .....	1
1.1.2 Objectives .....	2
1.2 Alzheimer’s disease as a multifactorial disorder .....	4
1.2.1 The role of amyloid.....	4
1.2.2 The role of tau.....	5
1.2.3 The role of vascular changes.....	6
1.2.4 The role of glucose metabolism.....	7
1.2.5 The role of neuronal activity.....	7
1.2.6 The role of atrophy.....	8
1.3 Parkinson’s disease as a multifactorial disorder .....	9
1.3.1 The role of dopamine .....	10
1.3.2 The role of white matter changes.....	11
1.4 Cellular vulnerability in Alzheimer's and Parkinson’s disease.....	12
1.5 Neurodegenerative disease progression modelling.....	13
1.5.1 Neuropathological staging .....	13

1.5.2	Hypothetical models .....	15
1.5.3	Mathematical models .....	16
1.6	Conclusion .....	18
Chapter 2 : Integrated transcriptomic and neuroimaging brain model decodes biological mechanisms in aging and Alzheimer’s disease.....		
		20
2.1	Preamble .....	20
2.2	Abstract.....	21
2.3	Introduction.....	21
2.4	Results.....	24
2.4.1	Capturing gene and macroscopic factor interactions in the human brain .....	24
2.4.2	Identifying genes driving biological and cognitive changes in healthy aging.....	27
2.4.3	Revealing top genes and molecular pathways controlling multifactorial alterations and clinical deterioration in AD .....	29
2.5	Discussion.....	31
2.5.1	Gene expression patterns modulate multifactorial interactions in healthy aging and AD progression .....	31
2.5.2	Aging and Alzheimer’s disease have both common and distinct mechanisms .....	34
2.5.3	Towards a genetic approach to extending healthy aging and treating Alzheimer’s disease .....	35
2.6	Materials and methods .....	37
2.6.1	Data description and processing .....	37
2.6.2	Gene Expression Multifactorial Causal Model.....	43
2.6.3	Model evaluation .....	44
2.7	Data availability .....	46
2.8	Supplementary Figures .....	47
2.9	Supplementary Tables.....	48
2.10	Article information.....	52
2.10.1	Acknowledgement .....	52
2.10.2	Ethics.....	53

Chapter 3 : Patient-centered transcriptomic and multimodal neuroimaging determinants of clinical progression, physical activity, and treatment needs in Parkinson’s disease .....	63
3.1 Preamble .....	63
3.2 Abstract.....	64
3.3 Introduction.....	64
3.4 Results.....	67
3.4.1 Whole-brain multiscale transcriptomic-neuroimaging model of Parkinson’s disease	67
3.4.2 Identifying transcriptomic mechanisms mediating behavioural and cognitive deterioration in PD .....	69
3.4.3 Uncovering the protein-protein interaction networks underlying PD phenotypic landscapes .....	71
3.4.4 Molecular pathways associated with physical activity in PD .....	73
3.4.5 Virtual Gene Perturbations Reveal Potentially Effective Drug Candidates .....	75
3.5 Discussion.....	77
3.6 Methods.....	83
3.6.1 Ethics statement .....	83
3.6.2 Data description and processing .....	83
3.6.3 Gene expression multifactorial causal model (GE-MCM) .....	89
3.7 Statistical analyses .....	90
3.7.1 Model fitting .....	90
3.7.2 Covariance of gene-neuroimaging parameters with clinical evaluations .....	90
3.7.3 Covariance of gene-neuroimaging parameters with physical activity .....	91
3.7.4 Gene perturbation for drug discovery .....	92
3.8 Supplementary Tables.....	93
3.9 Data and code availability.....	93
3.10 Acknowledgements.....	93
Chapter 4 : Single-nucleus RNA velocity reveals critical synaptic and cell-cycle dysregulations in neuropathologically confirmed Alzheimer’s disease .....	103
4.1 Preamble .....	103
4.2 Abstract.....	104

4.3	Introduction.....	105
4.4	Results.....	106
4.4.1	Data origin and single-nucleus RNA velocity estimation.....	106
4.4.2	Static vs dynamic genetic-cell modifications underlying AD evolution .....	109
4.4.3	Several RNA-velocity differences underlie AD neuropathological severity.....	111
4.4.4	Cross-study validation of differential RNA velocity .....	113
4.5	Discussion.....	115
4.6	Methods.....	119
4.6.1	Dataset-1 (Prefrontal cortex).....	119
4.6.2	RNA abundance and cell type identification .....	120
4.6.3	RNA velocity estimation.....	121
4.6.4	Differential expression and RNA velocity analyses .....	121
4.6.5	Cell speed and residual velocity estimation.....	122
4.6.6	Correlation with neuropathology .....	122
4.6.7	Validation in independent dataset (Dataset-2: Dorsolateral prefrontal cortex) .....	122
4.6.8	Biological pathway analyses.....	123
4.6.9	Comparison between single-cell and single-nucleus RNA velocities (Dataset-3) ....	123
4.6.10	Data visualization.....	124
4.7	Supplementary Figures .....	124
4.8	Additional information.....	130
4.8.1	Data availability .....	130
4.8.2	Code availability .....	131
4.8.3	Acknowledgements.....	131
Chapter 5 : Discussion .....		137
5.1	Summary of findings.....	137
5.1.1	GE-MCM for aging and AD .....	137
5.1.2	GE-MCM for PD .....	138
5.1.3	Single-cell RNA velocity changes in AD .....	140
5.2	Limitations .....	141
5.2.1	GE-MCM for aging, AD and PD.....	141
5.2.2	RNA velocity changes in AD .....	143

5.3	Future direction.....	144
5.3.1	Integration of peripheral biomarkers .....	144
5.3.2	Incorporating cell-type specific changes .....	145
5.3.3	Analysis of comorbidity.....	145
5.3.4	Application of GE-MCM to other progressive diseases .....	146
5.3.5	Comparison between subtype and spectrum.....	146
5.3.6	Sex-specific analyses .....	147
5.3.7	Translation to clinical practice.....	147
5.4	Conclusion .....	147
	Bibliography .....	149

## List of Figures

Figure 1.1: Basal ganglia circuit of direct and indirect pathways of motor control in health and disease. ....	11
Figure 1.2 : Braak neuropathological staging for tau in Alzheimer's disease.....	14
Figure 1.3: Hypothetic models of Alzheimer's disease and Lewy body disease progression. ....	16
Figure 2.1: Modelling the gene-imaging interactions driving healthy aging and AD progression .....	25
Figure 2.2: Reconstruction of individual multifactorial alteration patterns across all subjects in the AD continuum.....	26
Figure 2.3: Identification of top genetic modulators of cognitive change in healthy aging. ....	28
Figure 2.4: Uncovering the top genetic determinants of AD progression. ....	30
Figure 2.S1:Subject selection.....	47
Figure 2.S2: Correlation of predicted mRNA expression with actual mRNA expression across 976 genes. ....	48
Figure 3.1: Transcriptomic-neuroimaging multifactorial causal modeling of PD.....	68
Figure 3.2: Genes underlying clinically-relevant spatiotemporal vulnerability in PD overlap with cancer and infection pathways.....	70
Figure 3.3: Associations of patterns of PD clinical symptoms with biological mechanisms. ....	72
Figure 3.4 : Distinct molecular pathways are associated with leisure and work activities.....	75
Figure 3.5 : Top drugs for PD identified by our virtual perturbation framework.....	77
Table 3.S1: Distribution of number of time points of neuroimaging modalities.....	93
Table 3.S2: List of genes used in this study.....	93
Figure 4.1: Single-nucleus RNA-seq of the prefrontal cortex of 48 individuals across the Alzheimer's disease (AD) spectrum. ....	108
Figure 4.2: Differences in RNA velocity and gene expression underlying neuropathological AD progression. ....	111
Figure 4.3: Association of RNA velocity with relevant Alzheimer's-related neuropathological traits.....	112
Figure 4.4: Cross-study validation of RNA velocity differences in neuropathologic AD.....	114
Figure 4.S1: Validation of the correspondence between RNA velocities of single-cell and single-nucleus sequencing dataset from microglia of same subjects. ....	125
Figure 4.S2: Comparison of single-nucleus RNA velocity trajectory predictions between scVelo and veloVI.....	126
Figure 4.S4: Correlation between gene length and Wilcoxon rank sum test U-statistic (from differential velocity analysis between AD and control groups) in inhibitory neurons. ....	127
Figure 4.S5: Protein-protein interaction network of the differential velocity genes implicated in synaptic and ion channel activities in excitatory neurons.....	128

Figure 4.S6: Changes in RNA velocity with disease progression. .... 128

Figure 4.S7: Enriched disease pathway networks in male and female across six cell types. 129

Figure 4.S8: Percentage overlap of dynamic genes associated with AD pathology in each cell type..... 129

Figure 4.S9: Test for difference in post-mortem interval between the control and AD groups. .... 130

## List of Tables

Table 2.S1: Main demographic characteristics of the included ADNI subjects.....	48
Table 2.S2: List of 976 genes used in this study.....	49
Table 2.S3: Brain regions used in this study.....	49
Table 2.S4: Distribution of stable gene-imaging interaction parameters in healthy aging and AD progression (99% CI). ....	50
Table 2.S5: Identified molecular pathways underlying AD progression.....	50

## List of Abbreviations

A $\beta$	Amyloid $\beta$
AD	Alzheimer's Disease
ADNI	Alzheimer's Disease Neuroimaging Initiative
AHBA	Allen Human Brain Atlas
FA	Fractional Anisotropy
fMRI	functional Magnetic Resonance Imaging
GE-MCM	Gene Expression Multifactorial Causal Model
MD	Mean Diffusivity
MRI	Magnetic Resonance Imaging
PD	Parkinson's Disease
PET	Positron Emission Tomography
PPMI	Parkinson's Progression Marker Initiative
scRNA	single-cell RNA
snRNA	Single-nucleus RNA
SPECT	Single-Photon Emission Computed Tomography

## Acknowledgements

The work in this thesis spanned a period of almost six years of highs and lows. I wouldn't have completed it without the support of wonderful people, for whom if I pour my words of thanks into the Atlantic Ocean, the latter will overflow. I would like to thank my supervisor, Dr Yasser Iturria Medina, for his incredible mentorship. When I walked into MNI in 2019, his former lab members of his postdoc days said nothing about him, except goodness. It never took long before I confirmed all they said. Yasser supported me impeccably, eased my burden during financial difficulties, and taught me an important lesson that will rest with me for life. Whenever I made mistakes, he would say “the issue is not trivial, that was why you made mistake”. Deep inside me, I felt I could have been smarter. But Yasser's words were reassuring and positive. I keep this lesson with me and will always interact with others in a similar manner. Thank you, Yasser.

I was blessed with amazing advisory committee members, Drs Rosemary Bagot and Jorge Armony. I appreciate them for the guidance and support they provided throughout my PhD journey. Their selflessness, super-positive criticisms and feedback, their time, and commitment towards my PhD progress were sensational. I can't thank them enough.

I was surrounded by very supportive and humble lab colleagues. I particularly thank Ahmed Faraz Khan for being my discussion mate, especially when I was stuck on my project. His insights and contributions were piercing. I appreciate the postdocs and research associates, Drs Sue Lin Jin, Robert Baumeister, and Gleb Bezgin for their humility and willingness to share from their wealth of experience. I thank Lazaro Sanchez-Rodriguez, Veronika Pak, Mahdie Soltaninejad, Joon Hong, and Matthew Danyluik for their suggestions during lab meetings and for making the lab environment convivial.

I would like to thank my mother, Folake Adewale, and father, Waheed Adewale, for their toils and dedication of many years. I remember vividly when my father would say to

himself “even if I would wear tattered clothes, the most important thing is my children’s education”. Dad, yours is a cloth of honor and fulfillment.

Importantly, I owe a great deal of my PhD success to my loving wife, Dr Roqeebat Olanrewaju. How would this PhD journey have been without you? I would have collapsed along the way and died of loneliness. Thank you for being my emotional fortress and unparalleled motivation. To my two pulchritudinous daughters, Rufaidah and Razeenah, I say thank you for bringing smiles to my face during the lows. Rufaidah would say, “*Abi, tu es tired?*” (meaning: “My dad, are you tired?”); her mixing of Arabic, French and English languages was amusing enough to lift my spirits. While I wrote this thesis, Razeenah would often jump in front of my computer and press my keyboard randomly. Praise be to Allah, by whose mercy all good things are accomplished.

*For those who taught me to write  
When I knew nothing of letters*

*For those who took my hands  
And held me up from fall*

*For those who helped me walk  
When I could barely stand*

*For those who showed me the way  
When I was almost lost*

*For those who walked with me  
Down the road of lonesomeness*

*For those who stayed with me  
In the umbra of a moonless desert*

*For those who showed me care  
When I had despaired of help*

*The birds chirp hymn of thanks  
And the trees sway in dance of praise*

## **Contribution of authors**

The original work in this thesis includes Chapters 2, 3 and 4. I am the sole first author of all the three manuscripts. My role involved study conceptualization and design, data processing and analysis, results interpretation, manuscript writing, and project management. However, all the manuscripts feature some co-authors who have supported the projects invaluablely, and their contributions are summarised below. All authors provided critical feedback that helped shape the research, analysis, and manuscripts.

### **Chapter 2: Integrated transcriptomic and neuroimaging brain model decodes biological mechanisms in aging and Alzheimer's disease**

- Ahmed Faraz Khan: Manuscript review and editing
- Felix Carbonell: Contribution: Statistical method, manuscript review and editing
- Yasser Iturria Medina: Conceptualization, manuscript editing and review, supervision
- Alzheimer's Disease Neuroimaging Initiative: Data generation

### **Chapter 3: Patient-centered transcriptomic and multimodal neuroimaging determinants of clinical progression, physical activity, and treatment needs in Parkinson's disease**

- Ahmad Faraz Khan: Data preprocessing, manuscript review and editing
- Sue-Jin Lin: Data preprocessing
- Tobias R. Baumeister: Statistical method
- Yashar Zeighami: Data preprocessing
- Felix Carbonell: Statistical method
- David Ferreira: Manuscript review and editing
- Yasser Iturria Medina: Conceptualization, manuscript editing and review, supervision

**Chapter 4: Single-nucleus RNA velocity reveals critical synaptic and cell-cycle dysregulations in neuropathologically confirmed Alzheimer's disease**

- Ahmed Faraz Khan: Data preprocessing, manuscript editing and review
- David A Bennett: Manuscript editing and review
- Yasser Iturria Medina: Conceptualization, manuscript editing and review, supervision

# Contributions to knowledge

## Chapter 2

- Development of a novel data-driven dynamical systems model that can be used to study progressive disorders.
- Integrating expression patterns of multiple genes and several neuroimaging modalities reveals the biological pathways underlying normal aging and Alzheimer's disease (AD).
- Interactions between gene expression and neuroimaging-derived biological measures uncovers the genes driving cognitive change in aging and AD.
- Confirming the functional roles of identified aging- and AD-associated genes, and identifying novel roles that can be validated experimentally.
- Normal aging and AD share some biological pathways (G protein-coupled receptor, apoptosis, oxidative stress, and immune response/inflammation), with AD having more dysregulated pathways.

## Chapter 3

- Integrating gene expression with multimodal neuroimaging shows that Parkinson's disease (PD) shares similar mechanisms with cancer and infections.
- Identification of five distinct PD symptom profiles which are associated with different biological pathways, protein-protein interaction networks and hub genes.
- Discovery of a novel hub gene for PD, and confirmation of five previously identified PD hub genes.
- Physical activity in PD is mediated by immune response and cholesterol biosynthesis.
- *In silico* perturbation identified putative PD drugs, including Levodopa which is the mainstay of PD treatment.

## Chapter 4

- Dynamical model of gene expression captured biological processes of AD that are not revealed by static gene expression patterns.
- Identification of synaptic and cell-cycle dysregulation as critical processes underlying AD pathogenesis across 6 distinct brain cell types (excitatory neurons, inhibitory neurons, astrocytes, microglia, oligodendrocytes, oligodendrocyte precursor cells).
- Changes in gene dynamics of AD patients is sex-dependent, with females exhibiting more broadscale changes.
- Accelerated cell cycle is observed in AD patients compared to controls, with the least acceleration observed in microglia.
- As AD progresses towards the later stage, more genes are dysregulated in microglia, astrocytes, and oligodendrocytes, implying a progressive immune dysregulation.

### Overall:

- Development and release of a novel data-driven framework for understanding complex disease processes and identifying putative drug targets.
- Immune response/inflammation is central to normal aging, AD and PD.

# Chapter 1: Introduction and Literature Review

## 1.1 Rationale and Objectives

### 1.1.1 Rationale

Over the last two centuries, the life expectancy of humans has increased steadily, thanks to healthcare innovations. However, increased lifespan is associated with greater disposition to neurodegenerative diseases (Guerreiro & Bras, 2015; Singh et al., 2019). Hence, as the world populations skews towards the elderly, the prevalence of age-related neurodegenerative disorders, such as Alzheimer's disease (AD), Parkinson's disease (PD), frontotemporal dementia, and amyotrophic lateral sclerosis (ALS) is increasing. The World Health Organization (WHO) also forecasts that neurodegenerative disorders will overtake cancer to become the second leading cause of death in developed countries by 2040 (Gammon, 2014; Heemels, 2016). AD and PD are the two most common forms of neurodegenerative disease. Unfortunately, the currently available non-controversial treatments are at best symptomatic. Understanding these two diseases and developing disease-modifying treatments have therefore become a pressing research need.

Reductionist approach has traditionally enhanced medical practice, making it easy to classify and understand diseases at the basic level. Consequently, the sporadic forms of many neurodegenerative disorders are characterized by the deposition of abnormal proteins (Jellinger, 2010; Jucker & Walker, 2013; Taylor et al., 2002). A typical case is AD whose pathophysiology is linked to the intracellular and extracellular depositions of  $\beta$ -amyloid and tau proteins, respectively. Similarly, PD is genetically and neuropathologically linked to the misfolding of presynaptic  $\alpha$ -synuclein protein. While this proteinopathy paradigm is relevant, the complexity and heterogeneity of these diseases are becoming acknowledged (Armstrong,

2020; D. M. Wilson et al., 2023). The diseases are now considered multifactorial, involving dysregulations in various biological processes that interact in a complex manner.

A major translational advancement in the last two decades is the identification of various biomarkers to study different disease-associated biological processes. Transcriptomics analysis have paved the way for deciphering the genetic underpinnings of neurodegeneration at the microscopic level (Dillman et al., 2017; Iturria-Medina et al., 2020; Johnson et al., 2020; Mostafavi et al., 2018; Tanaka et al., 2018). Specifically, single-cell RNA sequencing technologies enables the study of transcriptomic changes across different cell types (Conte et al., 2024). In parallel, molecular positron emission tomography (PET) scans, single-photon emission computed tomography (SPECT) scans, and magnetic resonance imaging (MRI) are facilitating the detailed characterization of macroscopic disease-related changes, such as dopaminergic loss, A $\beta$  deposition, tau accumulation, glucose hypometabolism, altered cerebrovascular flow and atrophy (Booth et al., 2015; Dukart et al., 2013; Jack et al., 2018; Rodrigue et al., 2012; Zhang et al., 2017). Harnessing these multiple biomarkers may therefore hold clues to understanding how different processes interact in a complex manner to drive disease progression and heterogeneity in AD and PD.

### **1.1.2 Objectives**

The content of this thesis focuses on using dynamical systems analysis to integrate whole-brain gene expression with multimodal neuroimaging to characterize multiscale interactions between various biological processes that underlie AD and PD. The overall objective is driven by the critical need to advance the understanding of neurodegenerative diseases using available multiscale biological measures. This understanding will also facilitate the development of suitable computational tools for drug discovery and patient screening.

The objective of Chapter 1 is to discuss the motivation behind this thesis and its goals. Another objective is to review relevant literatures of AD and PD pathogenesis from a multifactorial perspective. The third objective is to critically examine the classes of models currently used in staging the two diseases and build a theoretical foundation for the data-driven models employed in this thesis.

The objective of Chapter 2 is to develop a data-driven multiscale model for studying disease progression. A second objective is to apply this model to understand multifactorial brain changes in normal aging and AD by incorporating multiple disease-related neuroimaging and transcriptomic measures. The final objective is to investigate how these multifactorial changes are related to clinical symptoms, and then uncover the similarities and differences between normal aging and AD.

Disentangling the heterogeneity of PD remains a research priority. Hence, the primary objective of Chapter 3 is to understand the multifactorial changes underlying PD and how those changes engender diverse clinical symptoms. A secondary objective is to identify the biological substrates of physical activity in PD. The last objective is to test the clinical utility of the developed data-driven model for drug discovery.

Neurons are notably affected in AD, but research suggests the roles of other brain cell types in the disease's development. The objective of Chapter 4 is to investigate the dynamic molecular changes occurring to the AD brain across various cell types and describe the biological processes driving these changes.

Lastly, the objective of Chapter 5 is to summarize the main findings in this thesis and discuss their implication for a better understanding of normal aging and neurodegeneration, and the development of computational tools for drug discovery.

## 1.2 Alzheimer's disease as a multifactorial disorder

Alzheimer's disease was first discovered in 1906 by Alois Alzheimer after carrying out an autopsy on Auguste Deter, a patient who had suffered from an unusual form of dementia antemortem (Stelzmann et al., 1995). Alois found the deposition of two peculiar substances in the cortex of this patient, and the substances were later identified as amyloid plaques and neurofibrillary tangles. Alois termed the disease presenile (early-onset) dementia because the patient developed the disease before 65 years. Indeed, some 10-15% of early-onset forms of AD are familial, and they have strong causal link with the two hallmark proteins (Ayodele et al., 2021). However, the sporadic forms of both early- and late-onset AD display a much complex biology that suggests the involvement of multiple mechanisms in the disease's aetiology.

### 1.2.1 The role of amyloid

The first major hallmark of AD pathology is the deposition of extracellular amyloid. Following the observation of a pathogenic mutation in amyloid precursor protein (*APP*) gene on chromosome 21, John Hardy and colleagues formulated the amyloid cascade hypothesis which posits that the aggregation and deposition of amyloid causes neurodegeneration and cognitive impairment in AD (J. A. Hardy & Higgins, 1992; J. Hardy & Allsop, 1991). In familial AD, mutations in genes that encode presenilin 1 (*PSEN1*), presenilin (*PSEN2*) or amyloid precursor protein (*APP*) disrupts the metabolism of  $\beta$ -amyloid (Borchelt et al., 1996; Citron et al., 1997; Scheuner et al., 1996). These genes display perfect penetrance as the mutation carriers almost certainly develop cognitive problems. However, familial AD accounts for less than 1% of AD cases, with the majority being sporadic (Bekris et al., 2010). The failure or minimal efficacy of many anti-amyloid therapies have further raised questions on the simplicity of the causal relationship between amyloid and neurodegeneration (Kim et al., 2022). As a result, the amyloid cascade hypothesis has faced criticisms over the years, and

there are constant calls for the incorporation of other factors with likely causal roles (Alawode et al., 2022; Morris et al., 2014).

Over the last two decades, a large body of evidence has pointed to influence of many risk factors for developing AD. More than 29 risk genes have been identified including *APOE*, *ABCA7*, *GAB2*, and *PICALM* (Bertram & Tanzi, 2019). Consequently, a probabilistic model of the amyloid cascade hypothesis was recently proposed to account for genetic risks (Frisoni et al., 2021). The authors posit three variants of AD, namely autosomal dominant AD, sporadic AD with *APOE*  $\epsilon$ 4 penetrance, and sporadic AD without *APOE* influence. However, this probabilistic model still identifies amyloid as the main causal driver of AD.

### **1.2.2 The role of tau**

The second hallmark of AD pathology is the presence of intracellular neurofibrillary tangles made of hyperphosphorylated tau. Tau is a microtubule associated protein (*MAPT*) and it is highly expressed in neurons (Terwel et al., 2002). The primary function of tau is to stabilize microtubules, which are cytoskeletal filaments that maintain neuronal structure organization and aid axonal transport of organelles such as lipids, proteins, and synaptic vesicles (Muralidar et al., 2020). Tau comprises four domains, namely N-terminal region, proline-rich region, microtubule-binding repeat region, and C-terminal region (Y. Chen & Yu, 2023). To stabilise microtubules, tau's microtubule-binding repeat region binds to the interior of microtubules while the proline-rich region binds to the surface of microtubules. Alternative splicing of tau's *MAPT* gene yields isoforms comprising of 3 or 4 binding repeat regions (3R or 4R). The tau aggregates in AD is a mixture of 3R and 4R, whereas other tauopathies such as progressive supranuclear palsy, corticobasal degeneration and Pick's disease contain only 3R or 4R aggregates (Buchholz & Zempel, 2024).

Tau undergoes several post-translational modifications including phosphorylation, acetylation, methylation, and O-glycosylation (Congdon & Sigurdsson, 2018).

Phosphorylation is the most studied because it is found in many tauopathies and can independently cause tau aggregation. Tau is hyperphosphorylated by about 4 folds in AD brains compared to controls (Köpke et al., 1993). While phosphorylation of tau enhances neuronal plasticity, hyperphosphorylation weakens the interaction of tau with microtubules, thereby facilitating tau dissociation and aggregation. The aggregated tau is believed to spread in a prion-like and stereotypical manner from the entorhinal cortex to the hippocampus and eventually to the entire cortex (Braak & Braak, 1991; Braak & Del Tredici, 2015).

### **1.2.3 The role of vascular changes**

There is an important link between vascular function and AD (Cortes-Canteli & Iadecola, 2020; Dodge et al., 2017). In fact, a data-driven study suggested that vascular dysregulation may precede amyloid deposition in AD pathogenesis (Iturria-Medina et al., 2016). Signs of vascular pathology are observed in more than 50% of AD cases (Sweeney et al., 2019). Both AD and vascular dementia have overlapping symptoms, even though the latter is defined as a clinical entity where patients must present neurocognitive problem and aetiology consistent with vascular cerebrovascular event or impaired attention and frontal-executive function. The definition of pure AD and pure vascular dementia remains controversial, and either of the diseases rarely occur in isolation (Groves et al., 2000; Jellinger & Attems, 2010).

A major vascular problem that has been consistently associated with AD is reduced blood flow (hypoperfusion) (Wolters et al., 2017). Accelerated cognitive decline is observed in AD patients having hypoperfusion (Duncombe et al., 2017). Furthermore, white matter hyperintensity which is observed in AD and other neurological conditions is caused (among other factors) by reduced blood flow (Garnier-Crussard et al., 2023; C. J. Huang et al., 2022). Hypoperfusion can trigger many neurodegenerative pathological cascades such as hypoxia-induced oxidative stress, mitochondrial dysfunction, and inflammation (Inoue et al., 2023).

#### **1.2.4 The role of glucose metabolism**

The brain, like other body parts, requires adequate energy for proper functioning. Although weighing a meagre 2% of the total body mass, the resting brain consumes about 20% of the total body glucose-derived energy (Rink & Khanna, 2011). In AD, dysregulation in glucose metabolism is detected at various stages in a region-specific manner. At the early stage, hypometabolism is found in the anterior cingulate cortex and hippocampus (P. Chen et al., 2021; Ferrari et al., 2019). Later disease stage involves hypometabolism in frontal, occipital, and parietal regions (Ferrari et al., 2019). Impaired glucose uptake may even precede amyloid plaque deposition (J. Huang et al., 2024).

Glucose dysregulation is triggered in several ways. One way is the impaired action of glucose transporters (GLUT) which are responsible for moving glucose into the brain via the vasculature. Reduced expressions of *GLUT1* and *GLUT3* at the blood brain barrier have been shown to reduce glucose uptake by the brain (J. Huang et al., 2024). Another causal mechanism of glucose hypometabolism is insulin dysregulation. Insulin receptors are expressed throughout the brain and the highest densities are found in regions mainly implicated in AD such as the entorhinal cortex and hippocampus (Sedzikowska & Szablewski, 2021). Whenever glucose is needed by a cell, insulin signals the cell to obtain glucose via the blood stream. Insulin also facilitates the movement of glucose transporters across membranes. The movement of the insulin-regulated transporter *GLUT4* into hippocampal neurons affects spatial memory, which is consistent with high metabolic demands during cognitive tasks (McNay et al., 2010; Pearson-Leary & McNay, 2016).

#### **1.2.5 The role of neuronal activity**

While at rest, neurons in the brain fire spontaneously and continuously, forming the basis of brain circuit and computation (Newbold et al., 2020; Papadimitriou et al., 2020). This spontaneous activity depends on intrinsic excitability of the neurons as determined by their

connections and synaptic characteristics, incoming signals from other connected brain regions, and functional connectivity (Radulescu et al., 2023). One theory posits that the spontaneous neuronal activity is related to learning and memory, as demonstrated in learning replay in the hippocampus for memory consolidation during slow wave sleep (Buzsáki, 1989; Cannon & Miller, 2016; Ólafsdóttir et al., 2018). This explains why aberrant neuronal activities are strongly associated with cognitive and memory impairments.

Spontaneous neuronal activity is altered in some brain regions of AD patients (Liu et al., 2014). In fact, the default mode network comprising brain regions that are mostly active during wakeful rest have been particularly affected. These regions include the medial temporal lobe, posterior cingulate cortex, medial prefrontal cortex, temporoparietal areas and the precuneus, most of which experience a high load of amyloid and tau pathology in AD (Buckner et al., 2005; Jacobs et al., 2013). Furthermore, landmark studies of human and animal models of AD suggested that dysregulation of spontaneous neuronal activity is an early sign of AD (Hall et al., 2015; Šišková et al., 2014). In CA1 pyramidal neurons of the hippocampus of AD mouse, intrinsic hyperexcitability was observed before amyloid plaque deposition (Busche et al., 2012). However, it appears that hyperactivity occurs at early disease stage while hypoactivity occurs later as the disease progresses (Dickerson et al., 2005; O'Brien et al., 2010).

### **1.2.6 The role of atrophy**

A principal hallmark of almost all neurodegenerative diseases is neuronal loss. Atrophy is used to describe the loss of neuronal cells and their connections which results in morphological changes and reduction in brain volume. Atrophy is commonly regarded as part of normal aging, but the changes are more pronounced in neurodegenerative diseases (Blinkouskaya et al., 2021; Kesidou et al., 2023). Neurons are susceptible to atrophy because their axons and dendrites run over long distances, making the requirement for structural

maintenance particularly costly. Also, neurogenesis in the adult human brain is limited, hence the damages caused by aging are not easily replaced (D. M. Wilson et al., 2023) .

Atrophy is observed at various stages of AD in a region-specific manner. The pattern of atrophy in the neocortex closely recapitulates the pattern of tau distribution, especially at the later disease stage (Joie et al., 2020; Planche et al., 2022; Thompson et al., 2003). Neuronal atrophy starts from the hippocampus and entorhinal cortex, spreads to the temporal, parietal and frontal areas, and then to the motor areas (Whitwell, 2010). Hippocampal atrophy correlates with memory deficits in AD patients (Gosche et al., 2002; Jack et al., 2002; McDonald et al., 2012). Thus, it is considered one of the most validated and consistent biomarkers of AD, even though it is observed in other forms of dementia (Jack et al., 2011).

### **1.3 Parkinson's disease as a multifactorial disorder**

In 1817, Dr James Parkinson observed physical features of 6 six patients in his monograph – An Essay of the Shaking Palsy. In what he termed as *paralysis agitans*, patients presented predominantly motor symptoms such as tremor, slowness of movement (bradykinesia), rigidity and postural imbalance (Parkinson, 2002). It was in 1912 that Dr Friedrich Heinrich Lewy found clumps of proteins in brains of people who had died of PD symptoms. This protein was later called Lewy bodies and found to contain misfolded  $\alpha$ -synuclein (Holdorff, 2002).  $\alpha$ -synuclein in the midbrain and cortex has since then become a pathological hallmark of PD (Dijkstra et al., 2014; Olanow & Brundin, 2013). The second pathophysiological hallmark of PD is the depletion of dopamine in the midbrain (Cramb et al., 2023; Ramesh et al., 2023). Mutations in specific genes such *SNCA* (encoding  $\alpha$ -synuclein), *LRRK2*, *PINK1*, and *PRKN* are found in familial forms of PD. However, about 85% - 90% of PD cases are sporadic (Schulze et al., 2018; Tran et al., 2020). Similar to AD, atrophy and neuronal activity dysfunction are observed in PD (Heo et al., 2020; Zeighami et al., 2015). Other mechanisms include oxidative stress, mitochondrial dysfunction, and excitotoxicity

(Dong-Chen et al., 2023; Maiti et al., 2017). Nevertheless, the subsequent subsections will focus on the biological factors that were not previously discussed under AD or whose biomarkers are used in this thesis.

### **1.3.1 The role of dopamine**

Dopamine is a neurotransmitter produced in several parts of the brain including the basal ganglia and ventral tegmental area (Juárez Olgún et al., 2016). The basal ganglia are comprised of a group of nuclei that control motor function. These nuclei include the striatum, globus pallidus, substantia nigra, and subthalamic nucleus. It is believed that loss of dopaminergic neurons in the substantia nigra pars compacta are responsible for motor symptoms in PD. Indeed, dopamine deficit is observed at early disease stage and about 80% of dopaminergic neurons are lost at the time of PD diagnosis (Halliday & McCann, 2010). Levodopa, the mainstay of PD treatment, alleviates PD motor symptoms by converting to dopamine (Mi et al., 2007).

The dopamine-dependent circuit in the basal ganglia is dysregulated in PD (McGregor & Nelson, 2019). This circuit consists of two pathways for motor control, namely direct and indirect pathways (Figure 1.1). The spiny projection neurons of the striatum express D1 and D2 types of dopamine receptors, which are associated with the direct and indirect pathways, respectively. The direct pathway promotes movement while the indirect pathway opposes movement, suggesting that dopamine has both excitatory and inhibitory effects. Hyperactivity of the direct pathway and hypoactivity of the indirect pathway are believed to underlie motor impairments in PD (Gerfen et al., 1990; Obeso & Lanciego, 2011). Also, this circuit dysregulation might be responsible for the abnormal neuronal activity observed in PD patients (McGregor & Nelson, 2019). However, some human studies of PD suggested that abnormal neuronal activity could emanate from downstream of the striatum or even outside the basal ganglia (Gerfen & Surmeier, 2011; Valsky et al., 2020).

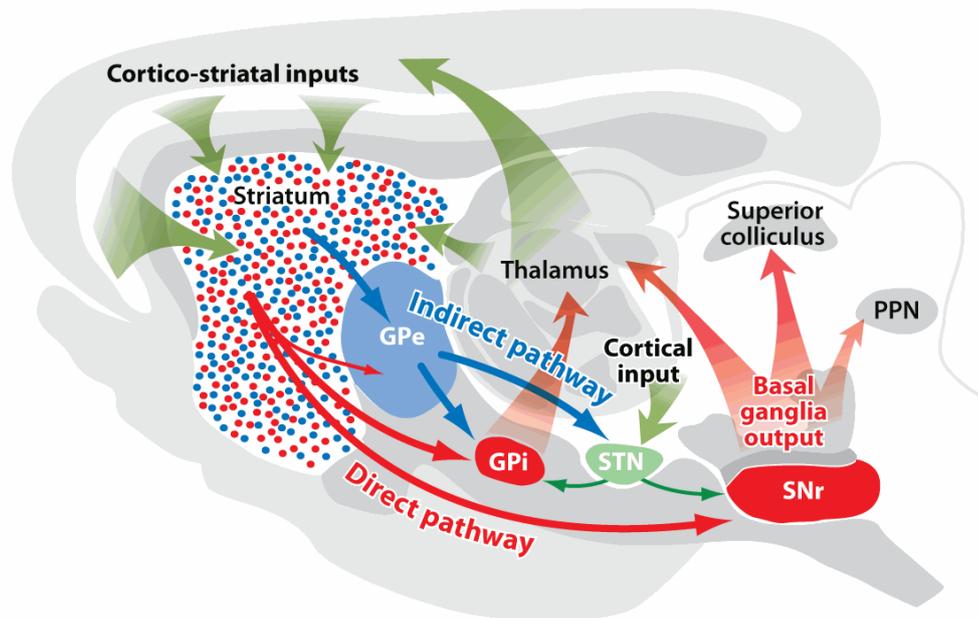


Figure 1.1: Basal ganglia circuit of direct and indirect pathways of motor control in health and disease. Original figure reproduced with permission from (Gerfen & Surmeier, 2011).

### 1.3.2 The role of white matter changes

The cell body and dendrite of neurons form the brain's grey matter while their myelinated axons make up the white matter. Due to the role of myelination in efficient transmission of signals across neurons, loss of myelin or axons could impair normal brain function (Baumann & Pham-Dinh, 2001). Although PD is primarily considered a disease of grey matter, white matter alterations are observed even before grey matter atrophy (Rektor et al., 2018). It is hypothesized that these white matter changes may represent an adaptive compensatory mechanism of the brain to preserve normal motor and cognitive functions.

Typical white matter changes in PD include alterations to the structure and composition of oligodendrocytes that form the myelinating cells of axons (Annese et al., 2013). Glial cytoplasmic inclusions is observed in the oligodendrocytes of PD patients carrying  $\alpha$ -synuclein mutations (Pasanen et al., 2014). Other neuroimaging studies identified white matter hyperintensities and microstructural changes in PD (Patriat et al., 2022; Vercruyssen et al., 2015). Fractional anisotropy (FA) is a diffusion tensor imaging (DTI) measure that detects

axonal or myelin alterations through directional constraint of transmission of water molecules within the axon. Mean diffusivity is another DTI measure that captures overall cellular changes of the axon. Motor symptom severity is correlated with lower FA and higher MD in the nucleus basalis of Meynert of PD patients (Nazmuddin et al., 2021). Other studies pointed to damage of white matter tracts in PD patients with cognitive symptoms (Agosta et al., 2014; Nakanishi et al., 2012).

## **1.4 Cellular vulnerability in Alzheimer's and Parkinson's disease**

The presentation of diverse clinical symptoms spanning several domains (e.g., cognitive, psychiatric, motor) in AD and PD patients suggests that certain brain regions and cell types might be preferentially vulnerable to pathological processes. The pattern of spread of atrophy and proteinaceous deposits in these neurodegenerative diseases also supports region- and cell-type-specific vulnerability to pathology (Braak et al., 2003; Braak & Braak, 1991; Cho et al., 2016). Thus, the traditional Braak stagings of AD and PD provide a sequential ordering of selective vulnerability. Beyond broad-class cell types (e.g., neuron versus glia) (Zimmer et al., 2024), different classes of a particular cell type might be differently susceptible to disease process or death because of triggers (Lee et al., 2021). These triggers could be unknown as in the case of sporadic forms of diseases or due to mutation as in the familial forms (Kampmann, 2024). Brain cells are selectively vulnerable via either cell autonomous or non-cell autonomous mechanisms (Gonzalez-Rodriguez et al., 2020; Wang et al., 2020). Cell autonomous mechanisms involve processes that affect a particular cell irrespective of the surrounding cell type or tissue. Some cell-autonomous mechanisms include genetic effects, cytoarchitecture, and cell type. Non-cell autonomous mechanisms include vascular or anatomical network, and cell-cell communication. Hence, vulnerability is both driven innately within the cell and by the spread of pathological proteins via cell-cell connections and communication.

Transcriptomic analysis of neurotypical brain may reveal the innate vulnerability of a cell to pathological stress (Sepulcre et al., 2018). Regional expression patterns of *MAPT* are associated with neuronal dysregulation in PD (Rittman et al., 2016). In AD, regions that are vulnerable to amyloid plaque showed under-expression of certain genes involved in protein synthesis and mitochondrial functions (Grothe et al., 2018). Conversely, regions that are vulnerable to atrophy displayed under-expression of genes involved in neuronal plasticity. However, due to the complexities in tissue processing and other technical limitations, most of the foregoing studies have focused on specific brain regions. Brain-wide gene expression analysis can better elucidate the influence of cell-autonomous processes on neurodegenerative cascades. Single-cell sequencing technologies also provides an unprecedented opportunity to study the genes and processes driving selective regional vulnerability in glial cells.

## **1.5 Neurodegenerative disease progression modelling**

### **1.5.1 Neuropathological staging**

Traditionally, neurodegenerative disease progression is assessed by neuropathological studies of post-mortem brain tissues, based on the frequency of affected regions across individuals (Braak et al., 2003; Braak & Braak, 1991; Josephs et al., 2014; Thal et al., 2002). The extent of disease progression is often indicated by the topography of abnormal protein deposition and neuronal loss. Pioneering works by Braak and Thal staged AD into five and six phases based on  $\beta$ -amyloid and tau depositions, respectively (Braak & Braak, 1991; Thal et al., 2002). As shown in Figure 1.2, the six Braak stages are defined by tau inclusions in the transentorhinal and entorhinal cortex (stage 1-II), subcortical limbic regions (stage III-IV),

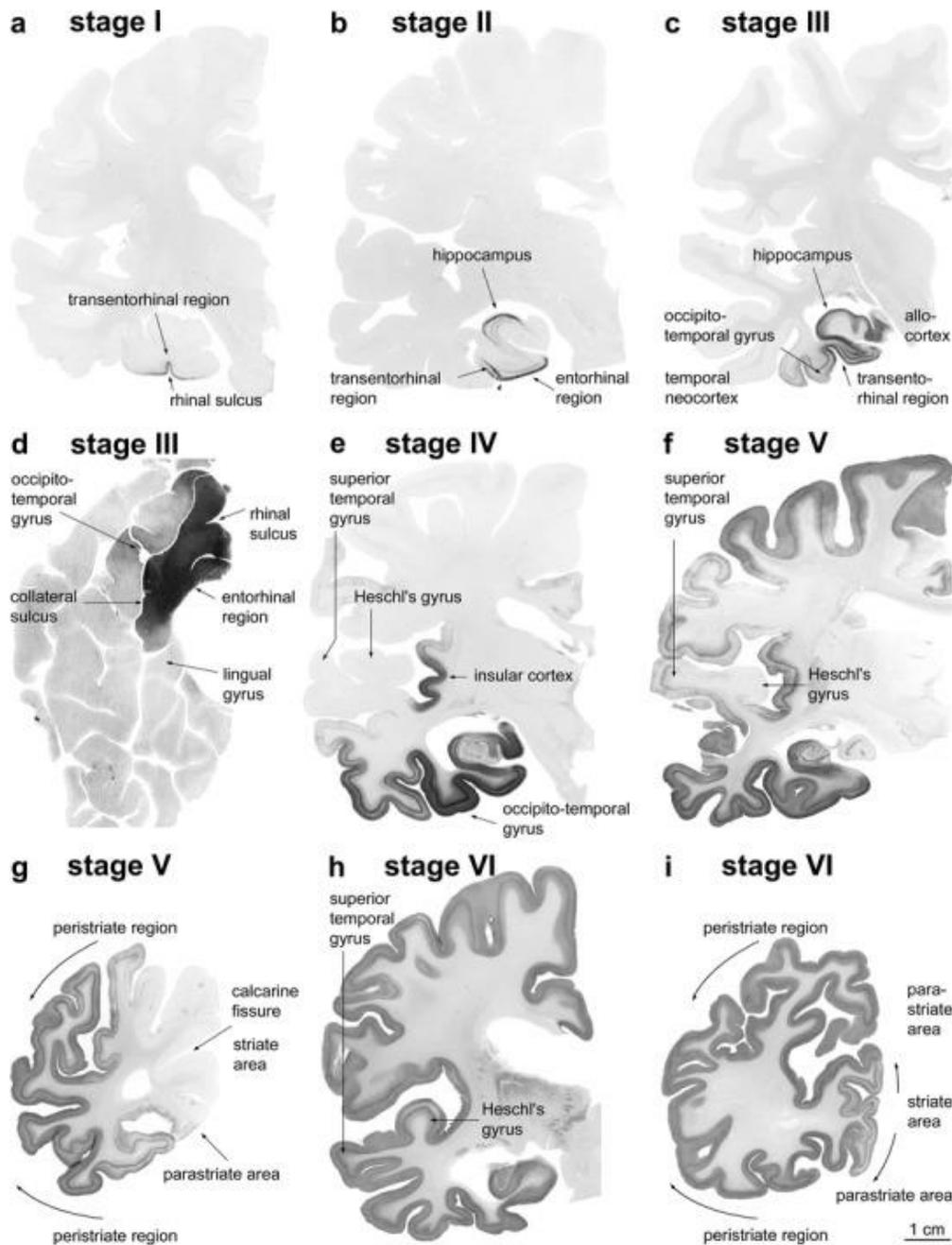


Figure 1.2 : Braak neuropathological staging for tau in Alzheimer's disease. Original figure reproduced from (Braak & Braak, 1996).

and the neocortex (stage V-VI). These stages of  $\beta$ -amyloid deposition is comprised of phase 1 for major parts of the neocortex, phase 2 involving the limbic areas, phase 3 for subcortical regions especially the basal ganglia, phase 4 for midbrain and basal forebrain, and phase 5 for pons and cerebellum. Braak and colleagues also proposed six stages for PD, based on the regions affected by Lewy bodies as disease progresses (Braak et al., 2003). The stages include

1-II for brainstem and olfactory bulb, II-IV for the midbrain and allocortex, and V-IV from thalamus to isocortex.

Disease staging based on post-mortem pathological studies neglects heterogeneity by assuming that neuropathology progresses stereotypically across regions in all subjects. In addition, it doesn't account for comorbidity, despite that more than 60% in older adults display co-pathology (Wennberg et al., 2019). Nevertheless, neuropathological staging forms the basis of other modelling methods discussed subsequently.

### **1.5.2 Hypothetical models**

Hypothetical models are used to suggest expected timelines of disease progression without incorporating much computation. Landmark study by Jack and colleagues drew upon evidence from *in vivo* and autopsy studies of AD to propose pathology ordering based on biomarker abnormalities (Jack et al., 2010). This hypothetical model assumes that biomarkers become abnormal sequentially and the abnormalities are associated with clinical symptoms. Using five well studied biomarkers (cerebrospinal fluid (CSF) A $\beta$ -42, CSF tau, fluorodeoxyglucose PET, amyloid PET, and structural MRI), 2 phases were hypothesised for AD progression. The early phase involves abnormality of amyloid biomarkers before the occurrence of neurodegeneration and clinical symptoms. The later phase involves abnormalities of biomarkers of neuronal dysfunction, neuronal injury (mediated by tau) and neurodegeneration. Similar hypothetical model has been proposed for Lewy body diseases wherein  $\alpha$ -synuclein deposition is followed by cell death and clinical symptoms (Donaghy & McKeith, 2014). As observable from Figure 1.3, hypothetical models generally base disease cascade on proteinopathy, implying that abnormal protein deposition is followed by cell damage and dysfunction, and then functional decline.

One of the main drawbacks of hypothetical models is their partial dependence on autopsy studies for biomarker ordering. Disentangling the sequence between tissue changes

and clinical symptoms at autopsy is challenging (R. S. Wilson et al., 2019). Hypothetical models have been continuously revised, with recent iterations accounting for individual differences in response to AD pathology, possibility of concurrent biomarker change, and the contributions of stochastic events such as genetics and environmental factors (Frisoni et al., 2021). However, there is still a need to incorporate inter-patient differences in terms of sequence of biomarker abnormality. Possible interactions between the different pathological processes should also be considered.

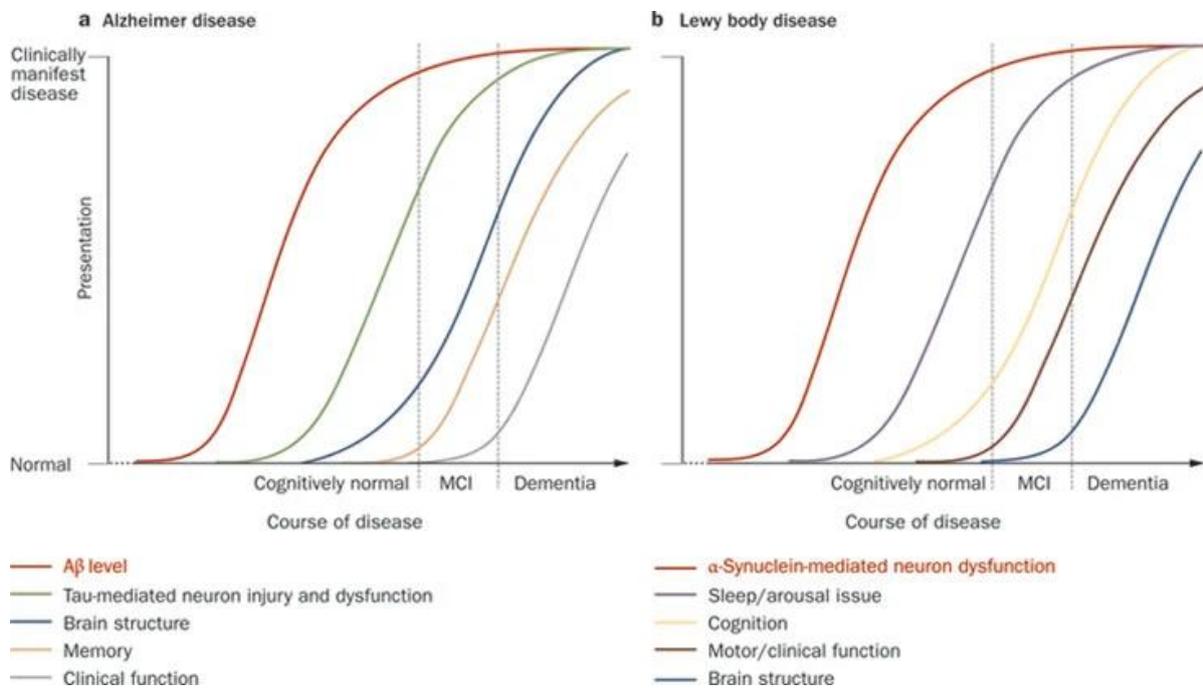


Figure 1.3: Hypothetic models of Alzheimer's disease and Lewy body disease progression. Original figure reproduced with permission from (Fields et al., 2011).

### 1.5.3 Mathematical models

Mathematical models use data-driven approach to characterize disease progression. They are grouped into empirical, semi-mechanistic and mechanistic models. Empirical models form the simplest form of data-driven modelling methods as they do not make any assumptions about the underlying biological process of the disease. Regression models generally fall under this category; they have been used to estimate response to treatment, and chart the timeline of

disease progression (Fleisher et al., 2015; Vu et al., 2012). In most regression models, the disease state at a particular time is defined as a function of the disease state at baseline. More advanced forms of regression such as linear mixed effects models are used to capture individual variabilities, where population level effects are fixed, and individual variabilities are captured as random effects representing deviation from population value (Guerrero et al., 2016). Although simple, empirical models do not account for biological constraints and may therefore miss important mechanistic processes. In addition, they do not account for possible interactions between different biomarkers.

Semi-mechanistic models involve more data-driven approach, and they also afford basic mechanistic explanation. Event-based models are one of the most common semi-mechanistic methods used in neurodegenerative disease modelling (Fonteiijn et al., 2012; Venkatraghavan et al., 2019; Young et al., 2014). By drawing inspiration from neuropathological staging, event-based models apply probabilistic methods to order biomarkers (or symptoms) as a sequence of irreversible events that transition from normal (healthy) to abnormal (disease) states. While foremost event-based models assumed specific event sequence, recent improvements have attempted to capture concurrent event (biomarker) changes (Tandon et al., 2023). Besides event-based modelling, deep generative learning have been used to jointly model the spatiotemporal ordering of biomarker changes (Abi Nader et al., 2020; Martí-Juan et al., 2023). Nevertheless, the biological insights from semi-mechanistic models are limited, and they still do not account for the possible interactions between disease-related processes.

Mechanistic models characterize disease progression by leveraging what is known about the pathophysiological mechanism. As many neurodegenerative diseases are believed to spread between brain areas, network spreading models are the foundation of mechanistic models of neurodegenerative diseases. The spread of disease between brain regions are posited

to result from selective vulnerability or prion-like propagation (Jucker & Walker, 2013; Leng et al., 2021). Pioneering network spread models used graph theory metrics to capture disease spread (Raj, 2021; Zhou et al., 2012). However, graph theory metrics produce static measures which may change as disease progresses. As a result, other studies employed dynamical systems models to better capture evolving disease processes mechanistically. Early dynamical system models include network diffusion models and epidemic spreading models, which have been used to recapitulate the pattern of production and propagation of  $\beta$ -amyloid and tau proteins (Iturria-Medina et al., 2014; Raj et al., 2012; Vogel et al., 2020). Most of these dynamical systems model considered one biomarker or disease process at a time. An effort to look at multiple biomarkers combined different topological profiles and graph metrics to mechanistically explain the appearance and spread of pathology (Garbarino et al., 2019).

A general limitation of most models we have considered up until now is that they did not consider the possible interactions between the various pathological process or biomarkers, e.g., influence of dopaminergic loss on neuronal activity. Lee and colleagues improved on this limitation by accounting for the influence of tau on  $\beta$ -amyloid (Lee et al., 2022). A multifactorial causal model (MCM) was also proposed by Iturria Medina and colleagues (Iturria-Medina et al., 2017). The MCM is a dynamical system model that accounts for mechanistic interactions between multiple disease-related processes. However, the processes considered by MCM are mainly macroscopic neuroimaging measures derived from PET and MRI. The influence of microscopic effects such as genetics or transcriptomics was not considered.

## **1.6 Conclusion**

Neurodegenerative diseases are complex and multifactorial disorders with heterogenous clinical presentation. Because current drugs only alleviate symptoms with limited success, there are more recent trends towards biological definitions of these diseases in a bid

to develop disease-modifying therapies (Höglinger et al., 2024; Simuni et al., 2024). Biomarkers are invaluable tools that can facilitate the understanding of disease biology. It is therefore important to harness as much available biomarkers as possible to puzzle out disease complexity. Mathematical models play important roles by incorporating biomarkers to chart disease progression timelines. Particularly, mechanistic models are used to confirm different hypotheses and draw insights from the underlying disease biology. Nevertheless, using multiple biomarkers in mechanistic models of disease progression is challenging. Hence, many mechanistic models only incorporate a few biomarkers at a single spatial resolution and do not assess causal mechanisms. This thesis seeks to utilize mechanistic models to study normal aging and neurodegenerative processes considering multiple biological factors and spatial resolutions.

# Chapter 2: Integrated transcriptomic and neuroimaging brain model decodes biological mechanisms in aging and Alzheimer's disease

## 2.1 Preamble

In this chapter, we develop a novel data-driven dynamical systems model called gene-expression multifactorial causal model (GE-MCM) which we applied to characterize the multifactorial changes in aging and AD. Using gene expression patterns of a thousand genes, six different neuroimaging modalities (capturing amyloid, tau, neuronal activity, atrophy, cerebral blood flow and glucose metabolism) and multiple clinical evaluations from longitudinal multicentre study, we identified genes underlying normal aging and AD. In addition, we revealed which of the six neuroimaging-derived biological processes are modified by these genes, providing mechanistic insights into the possible roles of the genes in healthy aging and disease. Furthermore, we unravelled the biological mechanism that differently and jointly drive normal aging and AD.

This work has been published as:

Adewale, Quadri, Ahmed F Khan, Felix Carbonell, and Yasser Iturria-Medina. 2021. "Integrated Transcriptomic and Neuroimaging Brain Model Decodes Biological Mechanisms in Aging and Alzheimer's Disease." *eLife* 10 (May). <https://doi.org/10.7554/eLife.62589>

## 2.2 Abstract

Both healthy aging and Alzheimer's disease (AD) are characterized by concurrent alterations in several biological factors. However, generative brain models of aging and AD are limited in incorporating the measures of these biological factors at different spatial resolutions. Here, we propose a personalized bottom-up spatiotemporal brain model that accounts for the direct interplay between hundreds of RNA transcripts and multiple macroscopic neuroimaging modalities (PET, MRI). In normal elderly and AD participants, the model identifies top genes modulating tau and amyloid- $\beta$  burdens, vascular flow, glucose metabolism, functional activity, and atrophy to drive cognitive decline. The results also revealed that AD and healthy aging share specific biological mechanisms, even though AD is a separate entity with considerably more altered pathways. Overall, this personalized model offers novel insights into the multiscale alterations in the elderly brain, with important implications for identifying effective genetic targets for extending healthy aging and treating AD progression.

## 2.3 Introduction

Innovations in healthcare and drug delivery have led to increase in human life expectancy. However, increased lifespan is accompanied by more predisposition to frailty and late-onset Alzheimer's disease (AD) (Guerreiro and Bras, 2015; Singh et al., 2019). Both healthy aging and AD are complex multifactorial processes, and understanding their molecular mechanisms is crucial for extending longevity and improving the quality of life (Alkadhi and Eriksen, 2011; Kowald and Kirkwood, 1996). Indeed, at the microscopic scale ( $\sim 10^{-6}$  m), transcriptomics and proteomics analyses of the brain have paved the way for deciphering the mechanistic underpinnings of healthy aging and AD (Dillman et al., 2017; Iturria-Medina et al., 2020; Johnson et al., 2020; Mostafavi et al., 2018; Tanaka et al., 2018). In parallel,

macroscopic ( $\sim 10\text{--}2\text{ m}$ ) imaging phenotypes from PET and MRI are facilitating the detailed characterization of brain changes, such as amyloid- $\beta$  ( $A\beta$ ) and tau accumulation, glucose hypometabolism, altered cerebrovascular flow, and atrophy (Dukart et al., 2013; Jack et al., 2018; Rodrigue et al., 2012; Zhang et al., 2017). However, in both aging and disease research, most studies incorporate brain measurements at either micro- (e.g., transcriptomics) or macroscopic scale (e.g., PET imaging), failing to detect the direct causal relationships between several biological factors at multiple spatial resolutions.

Although AD is characterized by the accumulation of amyloid plaques and neurofibrillary tangles, many other biological aberrations have been associated with the disease (neuroinflammation, vascular abnormalities, white matter hyperintensities), leading to changes in diagnostic criteria in recent times (DeTure and Dickson, 2019). The complexity of AD is further compounded by the interplay between these multiple biological factors. A growing body of evidence points to the synergistic interaction between  $A\beta$  and tau in driving neuronal loss, functional dysregulation, and glucose hypometabolism in AD (Iaccarino et al., 2017; Ittner and Götz, 2011; Pascoal et al., 2017; Pickett et al., 2019). Also, cerebral blood flow (CBF) promotes  $A\beta$  clearance, suggesting that vascular dysregulation could impact neuronal function and facilitate  $A\beta$  deposition (Qosa et al., 2014; Zlokovic, 2011). To account for the synergy between multiple biological factors, we previously introduced a multifactorial causal model (MCM) (Iturria-Medina et al., 2017), which uses multimodal imaging data to characterize the macroscale intra-regional interactions among any pair of biological factors (e.g., tau,  $A\beta$ , CBF) while accounting for the inter-regional spreading of the pathological alterations across axonal and/or vascular connections. However, this multifactorial model did not consider the microscopic properties of the modelled brain regions.

In an initial attempt to integrate brain variables at multiple scales, a few recent studies have used the regional expression patterns of pre-selected genes as complementary information

in intra-brain disease-spreading models (Freeze et al., 2018; Freeze et al., 2019; Zheng et al., 2019). Applied to Parkinson's disease (PD), improvements in the capacity to explain regional brain atrophy patterns were observed, based on each brain region's genetic predisposition to the disease. However, most of these studies have selected very specific genes already known for their crucial role in disease (e.g., *SNCA*, *TMEM175*, *GBA*), while disregarding the individual and combined roles of several other relevant gene candidates. Moreover, the analyses have focused on the influence of transcriptomics on a single biological factor at a time, without accounting for the multiplicity of biological alterations and interactions that occur at different spatial scales. As a result, we continue to lack brain generative models integrating a large set of genetic activities with multimodal brain properties.

An integrated multiscale and multifactorial brain model (from genes to neuroimaging and cognition) may be critical to further our understanding of both healthy aging and neurodegeneration, and engender the development of inclusive biomarkers for personalized diagnoses and treatment. Driven by this motivation, here we combine whole-brain transcriptomics, PET, and MRI in a comprehensive generative and personalized formulation, which we successfully validated in healthy aging and AD progression. This novel approach concurrently accounts for the direct influence of hundreds of genes on regional macroscopic multifactorial effects, the pathological spreading of the ensuing aberrations across axonal and vascular networks, and the resultant effects of these alterations on cognition. The proposed framework constitutes a promising technique for identifying effective genetic targets to prevent aging-related disorders and ameliorate existing neurodegenerative conditions.

## 2.4 Results

### 2.4.1 Capturing gene and macroscopic factor interactions in the human brain

Genes control many biological functions, and their dysregulation can cause abnormal development, accelerated aging, or disease (Kuintzle et al., 2017; Lee and Young, 2013). Aiming to characterize the direct influence of genes on multiple brain processes, here we have developed a multiscale and multifactorial spatiotemporal brain model (Figure 2.1A–C) linking whole-brain gene expression with multiple macroscopic factors typically quantified via molecular PET and MRI modalities (i.e., A $\beta$  and tau proteins, CBF, glucose metabolism, neuronal activity, and grey matter density). This novel approach, called Gene Expression Multifactorial Causal Model (GE-MCM; see 'Methods'), enables the quantification of gene-specific impacts on the longitudinal changes associated with each local macroscopic factor considered and gene-mediation effects on pairwise factor interactions (e.g., negative tau effects on neuronal activity) while accounting for the simultaneous spreading of the aberrant effects across physical brain networks (e.g., tau and A $\beta$  region-region propagation via anatomical and vascular connectomes). By using standardized gene expression (GE) maps (Hawrylycz et al., 2012), longitudinal multimodal imaging data, and a robust optimization algorithm, the GE-MCM identifies individual transcriptomic-imaging parameters controlling the dynamic changes observed in the macroscopic biological factors considered (Figure 2A–C). These personalized parameters are assumed to be the gene-specific deviations required for model fitting and, thus, they quantitatively measure individual gene dysregulation patterns. We hypothesized that the post-hoc analysis of these transcriptomic-imaging parameters will reveal essential pathogenetic mechanisms in health and disease.

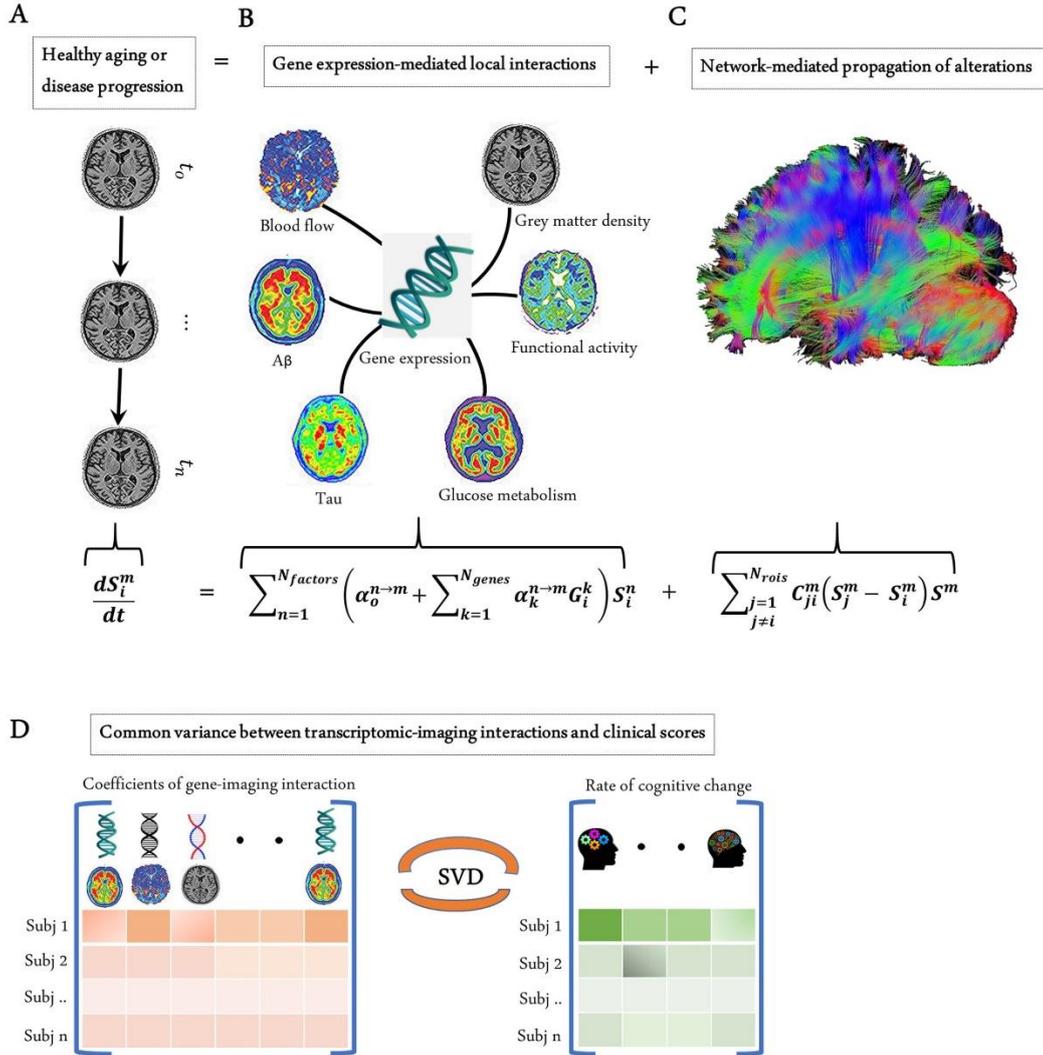


Figure 2.1: Modelling the gene-imaging interactions driving healthy aging and AD progression. (A) The longitudinal alteration of macroscopic biological factors in healthy and diseased brain due to gene-imaging interactions and the propagation of the ensuing alterations across brain network. (B) Regional multifactorial interactions between six macroscopic biological factors/imaging modalities are modulated by local gene expression. (C) Causal multifactorial propagation network capturing the interregional spread of biological factor alterations through physical connections. (D) By applying a multivariate analysis through singular value decomposition (SVD), the maximum cross-correlation between age-related changes in cognitive/clinical evaluation and the magnitude of genetic modulation of imaging modalities are determined in a cohort of stable healthy subjects (for healthy aging), mild cognitive impairment (MCI) converters, and Alzheimer’s disease (AD) subjects (for AD progression). The key causal genes driving healthy aging and AD progression are identified through their absolute contributions to the explained common variance between the gene-imaging interactions and cognitive scores.

Next, with the complementary interest of further clarifying the genetic mechanisms underlying healthy aging and AD development, the GE-MCM framework was applied to a cohort of 151 healthy and 309 diseased subjects from Alzheimer’s Disease Neuroimaging Initiative (ADNI) (see 'Methods' and Figure 2.1). The standardized transcriptomic data was

derived from six neurotypical brains from Allen Human Brain Atlas (AHBA) (Hawrylycz et al., 2012), comprising RNA intensities of 976 landmark genes with leading roles in central biological functions. These genes correspond to a set of universally informative transcripts, previously identified as ‘Landmark Genes’, based on their capacity to cover most of the information in the whole human transcriptome across a diversity of tissue types (Subramanian et al., 2017).

### Model performance across clinical groups

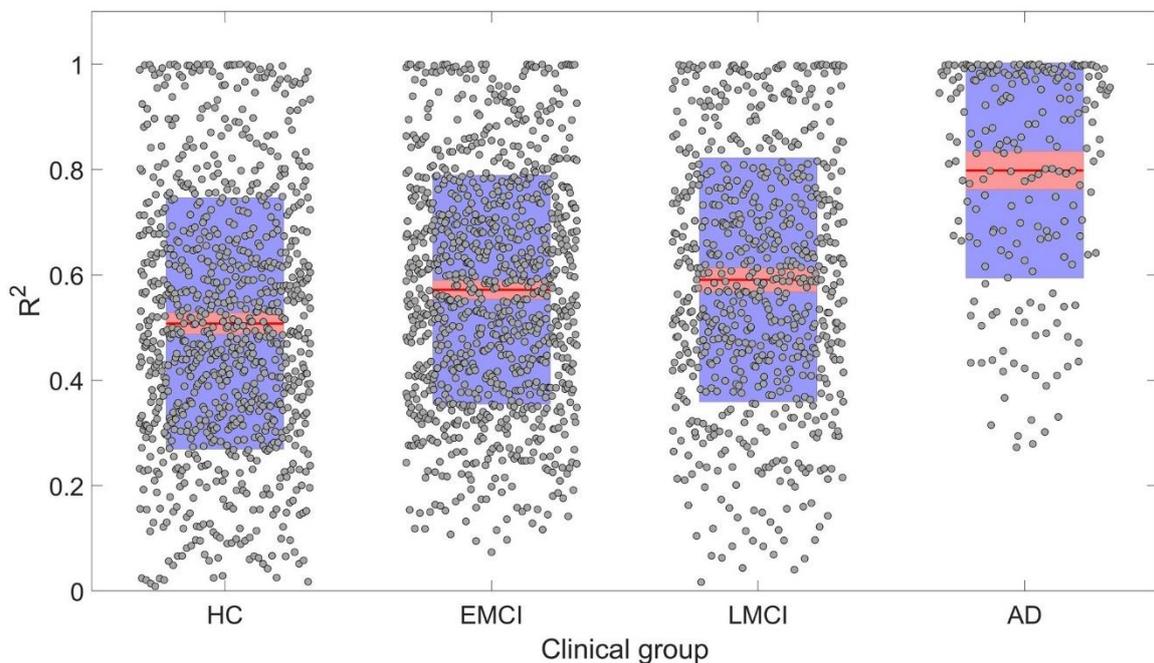


Figure 2.2: Reconstruction of individual multifactorial alteration patterns across all subjects in the AD continuum. Plots are shown for the  $R^2$  obtained across all six biological factors in the healthy control (HC) (n=151), early mild cognitive impairment (EMCI) (n=161), late mild cognitive impairment (LMCI) (n=113), and Alzheimer’s disease (AD) (n=35) cohorts. Points are laid over a 2.58 standard error of the mean (SEM) (99% confidence interval) in red and at 1 SD in blue. Notice that model performance improves with disease progression. We attribute this effect to the typical larger variation in longitudinal biological factor alterations with disease evolution, which provides the optimization algorithm with further biological information and results in a more accurate data fitting and parameter identification.

The predictive performance of the model across different clinical categories is shown in Figure 2.2. We calculated the coefficient of determination ( $R^2$ ) of the model for the six longitudinal PET and MRI modalities, and averaged them across all subjects in each clinical group. The  $R^2$  was highest for AD ( $0.80 \pm 0.20$ ), followed in order by late mild cognitive

impairment (LMCI) ( $0.59 \pm 0.23$ ), early mild cognitive impairment (EMCI) ( $0.57 \pm 0.21$ ), and healthy control (HC) ( $0.51 \pm 0.24$ ). The improvement observed in model performance with disease progression could be due to the larger variation in biological factor alteration in the later stages of the AD continuum. Nevertheless, these results support the capacity of the GE-MCM approach to reproduce the longitudinal observations in the six molecular PET and MRI modalities.

### **2.4.2 Identifying genes driving biological and cognitive changes in healthy aging**

Age is a significant risk factor for developing many complex disorders. Even though lifestyle and environmental factors contribute to healthy aging, understanding the genetic basis of aging will offer valuable biological insights with implications for disease prevention and longevity (Niccoli and Partridge, 2012; Rodríguez-Rodero et al., 2011). Hence, we sought to identify causal genes underlying longitudinal cognitive changes in healthy aging. We analysed the predictive relationship between the obtained transcriptomic-imaging parameters and multiple cognitive evaluations in 113 HC subjects who remained clinically stable within 7.8 years ( $SD = 2.9$  years). The cognitive changes were obtained as the age-related slopes of Mini-Mental State Examination (MMSE), Alzheimer's Disease Assessment Scale (ADAS), executive function (EF), and memory composite score (MEM) over an average of 7.2 time points ( $SD = 2.6$ ). For this analysis, we only used 68 stable transcriptomic-imaging parameters, the 99% CI of which excluded zero across the HC non-converters ('Model evaluation' subsection in 'Methods'). Using a multivariate singular value decomposition (SVD), we found the common latent variables between the gene-imaging parameters and the slopes of multiple cognitive measures, and the variances explained by the principal components (PCs) are shown in Figure 2.3A. Running 10,000 permutations, we identified the first PC as the only significant component (explained variance = 50.3%;  $p=0.0074$ ).

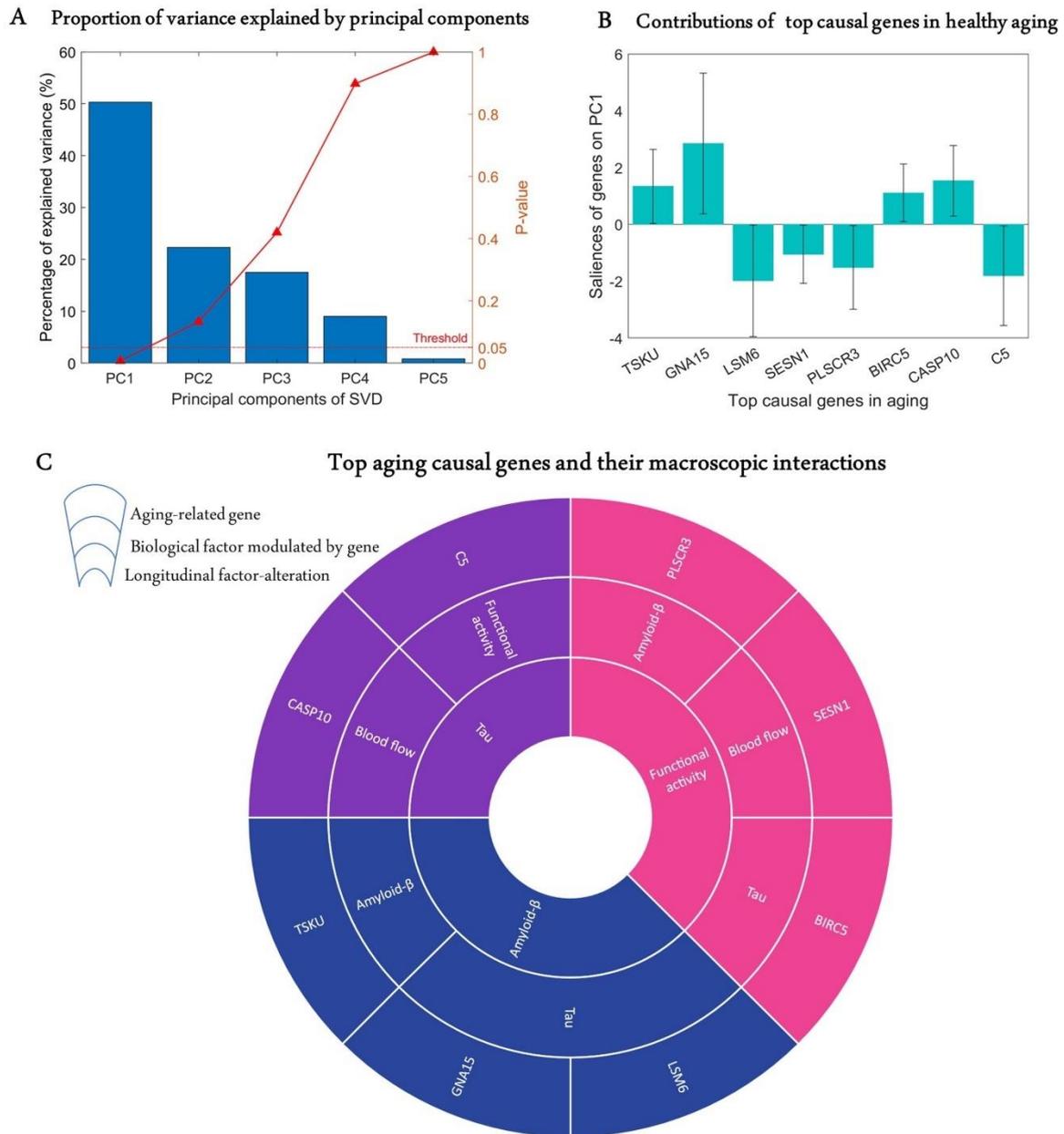


Figure 2.3: Identification of top genetic modulators of cognitive change in healthy aging. (A) Common variance (and associated p-values) captured by the top five principal components (PCs) of the singular value decomposition (SVD) in explaining the rate of change of cognitive scores due to healthy aging. Only the first PC is significant ( $p < 0.05$ ). (B) Genetic contributions (and 99% CI) on the first PC, depicted only for the eight highly stable aging-related genes, the bootstrap ratios of which are above 2.58. (C) Top genetic determinants of multifactorial alterations in healthy aging. The innermost ring shows the longitudinal biological factor altered with aging, the middle ring displays the interacting biological factors driving the longitudinal alteration, and the outermost ring represents the causal genes modulating the interactions among biological factors (e.g., SESN1 directly modulates blood flow to drive age-related alteration in neuronal activity).

Next, we calculated the contribution of each gene-specific parameter on this significant PC ('Model evaluation') and assessed the statistical reliability of the genetic contributions by running 10,000 bootstrap iterations. A bootstrap ratio threshold of 2.58 (approximately

equivalent to  $p < 0.01$ ; Efron and Tibshirani, 1986) was applied, revealing eight genes with stable causal contributions to the multimodal imaging dynamics and associated cognitive changes in healthy aging (Figure 2.3B). Notice that the saliences of some genes are negative, implying that their modulation effects are negatively associated with the rate of cognitive change. Specifically, as shown in Figure 2.3C, *TSKU* modulates  $A\beta$  while tau is modulated by *GNA15* and *LSM6* to drive age-related alterations in  $A\beta$ . Also, *BIRC5*, *SESNI*, and *PLSCR3*, respectively, modulate tau, CBF, and  $A\beta$  in driving alterations in neuronal activity. Similarly, age-related changes in tau are driven by *C5* and *CASP10* through their direct effects on functional activity and CBF, respectively.

### **2.4.3 Revealing top genes and molecular pathways controlling multifactorial alterations and clinical deterioration in AD**

A crucial challenge for the early detection and prevention of AD is the development of cheap and non-invasive biomarkers (such as genes) as well as the understanding of the molecular mechanisms underlying its pathogenesis (Iturria-Medina et al., 2020). Here, we proceed to identify genes driving neuropathological progression in the AD spectrum, restricting our analysis to 129 participants who were either diagnosed with AD (35) at baseline or converted to AD (94) after baseline diagnosis (7 HC and 87 MCI). Like the aging analysis, we only kept 993 statistically stable transcriptomic-imaging parameters, the 99% CI of which excluded zero ('Model evaluation' subsection in 'Methods'). We used SVD to obtain the common latent variables (variance) between the gene-imaging parameters and slopes of multiple cognitive measures (MMSE, ADAS, EF, and MEM across  $6.3 \pm 3.0$  longitudinal time points). After 10,000 permutation runs, the first PC was significant ( $p=0.009$ ) and explained 63.8% of the variance between the gene-imaging interaction parameters and the slopes of cognitive evaluations (Figure 2.4A). A bootstrap ratio threshold of 2.58 (approximately equivalent to  $p < 0.01$ ; Efron and Tibshirani, 1986) was applied, identifying 111 genes (Figure

2.4B) with stable causal contributions to the macroscopic factor interactions and associated cognitive changes in AD. The factors directly modulated by these causal genes and the ensuing factorial alterations are shown in Figure 2.4C.

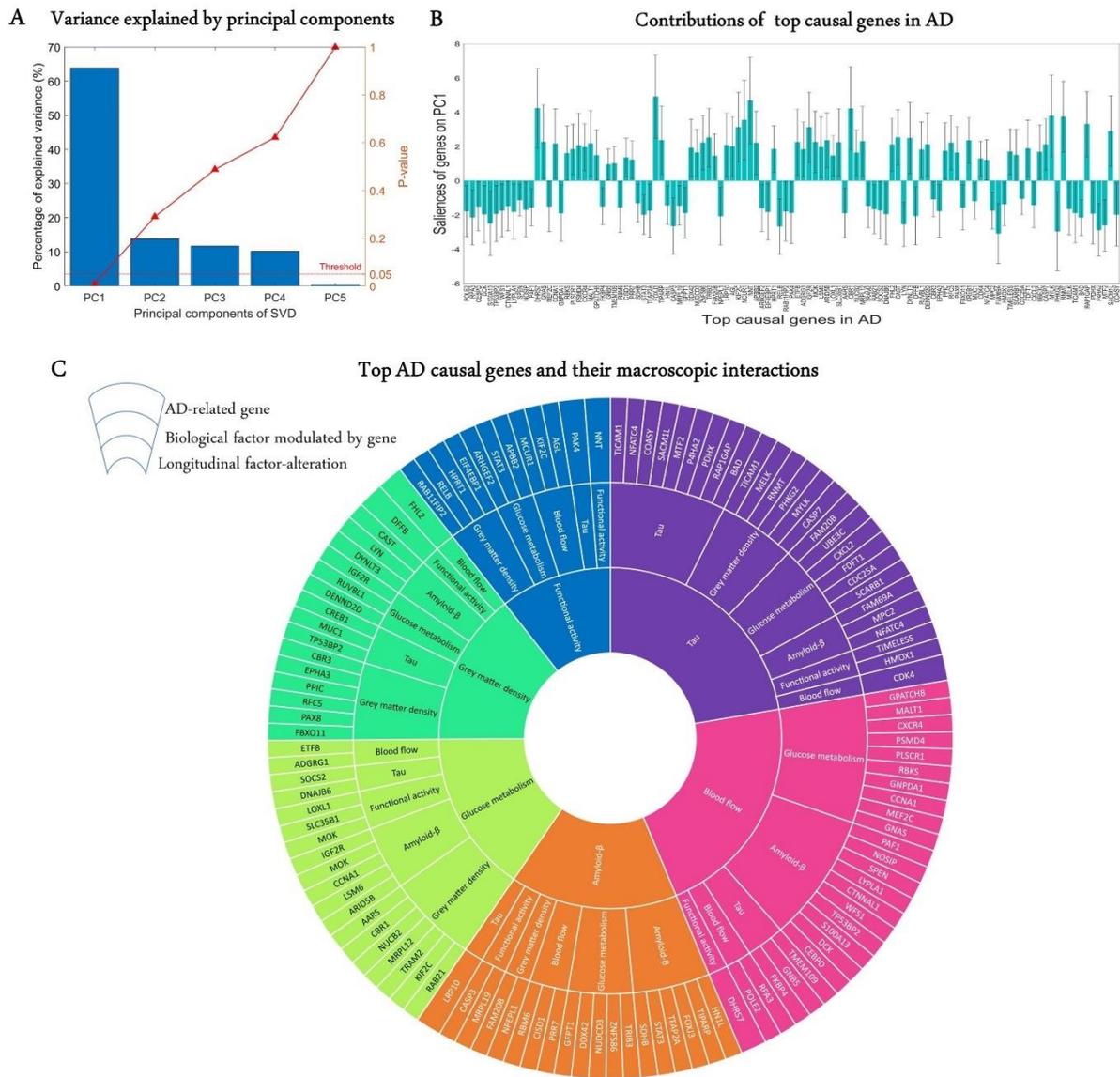


Figure 2.4: Uncovering the top genetic determinants of AD progression. (A) The common variance captured by the principal components (PCs) of the singular value decomposition (SVD) in explaining how clinical evaluations change with Alzheimer’s disease (AD) evolution. P-values after 10,000 permutations are also shown. (B) Contributions of top AD causal genes (with 99% CI) to the first PC. Top causal genes are identified by selecting those genes whose bootstrap ratios of saliencies are above 2.58. (C) Multifactorial interactions between the identified genes and imaging modalities. The innermost ring shows the longitudinal biological factor changes with AD, the middle ring displays the interacting biological factors driving the longitudinal alteration, and the outermost ring represents the causal genes modulating the interactions among biological factors. A gene directly influences how a biological factor interacts with other factors to cause a factorial alteration along the disease’s course.

Finally, we performed a large-scale gene functional analysis with PANTHER (Mi et al., 2013) to uncover the molecular pathways and biological functions associated with the 111 identified disease-driving genes. Sixty-five functional pathways were identified and most of them, including the Alzheimer disease-presenilin pathway, are highly representative of the biological processes commonly associated with neuropathology and cognitive decline (Table 2.S2). The pathways with the leading number of genes are apoptosis, cholecystokinin receptor signalling, inflammation mediated by chemokine and cytokine, and gonadotropin-releasing hormone receptor (see 'Discussion').

## **2.5 Discussion**

### **2.5.1 Gene expression patterns modulate multifactorial interactions in healthy aging and AD progression**

An unprecedented attribute of this study is the insight it provides into the multiscale interactions among aging and AD-associated biological factors, and the possible mechanistic roles of the identified genetic determinants. In concordance with our results in healthy aging (see Figure 2.3C), *BIRC5* have been shown to regulate microtubule dynamics and interact with tau (Zhao et al., 2003). Sestrins, including *SESNI*, preserve blood brain barrier integrity and serve a neuroprotective effect after cerebral ischemia (S.-D. Chen et al., 2019; Li et al., 2016; Shi et al., 2017) *C5* belongs to the complement immune system, and it modulates synaptic pruning and plasticity by interacting with microglia. (Wang et al., 2020).

Several animal and biostatistical studies also corroborate the functional relationships observed in AD results. In agreement with the interactions driving longitudinal alteration in blood flow (see Figure 2.4C), *FKBP4* encodes the *FKBP52* protein which has been demonstrated to alter tau phosphorylation pattern and stimulate its abnormal aggregation (Giustiniani et al., 2015). *FKBP52* also decreased significantly in brains of AD patients (Giustiniani et al., 2012). A bioinformatic and functional validation study identified the role of

*GNAS* in glucose metabolism through insulin regulation (Taneera et al., 2019). Notably, several GWAS and animal studies have consistently linked *MEF2C* to AD and its associated cognitive decline (Beecham et al., 2014; Davies et al., 2015; Lambert et al., 2013, (Sao et al., 2018)). Knocking out *MEF2C* in mice induced glucose metabolism impairment (Anderson et al., 2015). *PLSCR1* could drive atrophy due to its apoptotic effect and interaction with calcium ion in maintaining the organization of phospholipid bilayers of membranes (Sahu et al., 2007). *CXCR4* also regulates apoptosis and neuronal survival through glial signalling and the Rb/E2F pathway, respectively (Bezzi et al., 2001; Khan et al., 2008). Nitric oxide synthase interacting protein (*NOSIP*) controls the expression of nitric oxide synthase (NOS), the major source of nitric oxide in the brain (Dreyer et al., 2004). In brain endothelial cells, downregulating NOS upregulates *APP* (amyloid precursor protein) and *BACE1* ( $\beta$ -site APP-cleaving enzyme1) both of which control amyloid dynamics (Austin et al., 2010).

We also found congruous functional associations for the genes driving longitudinal alterations in  $A\beta$ . Apart from its apoptotic role, *CASP3* has been shown to regulate synaptic plasticity and functional activity *in vivo* (D'Amelio et al., 2010). *TRIB3* controls glucose metabolism, insulin signalling and the expression of other glucose metabolism genes ((Prudente et al., 2012; W. Zhang et al., 2013; W. Zhang et al., 2016). Among the genes altering tau with AD progression, nuclear factor of activated T cells (*NFAT*) overexpression in animal model increased  $A\beta$  production and promoted *BACE1* transcription (Mei et al., 2015). *TIMELESS* (*TIM*) is a gene with central role in controlling circadian neuronal activity (Kurien et al., 2019). Interestingly, dysregulated circadian rhythm is causally associated with AD (Homolak et al., 2018). Furthermore, our results on glucose metabolism dysregulation align with previous functional studies. *RAB21* may induce atrophy through apoptosis and cell growth inhibition (Ge et al., 2017). Due to its function in detoxifying reactive aldehydes produced from lipid peroxidation, the carbonyl reductase enzyme *CBR1* could prevent oxidative stress-

induced atrophy (Maser, 2006). *DNAJ* proteins belong to the group of chaperones that regulate protein homeostasis, and an earlier study implicated *DNAJB6* in  $\alpha$ -synuclein aggregation (Aprile et al., 2017). Investigating the effect *DNAJB6* on tau processing as suggested by our result could provide further insight into the roles of the gene in AD.

Supporting our results for longitudinal alterations in functional activity, downregulating *EIF4EBP1* prevents toxin-induced neuronal atrophy in PD model by blocking the action of apoptotic caspase-3 (Xu et al., 2014). The gene also mediates synaptic reorganization and refinement, independent of post synaptic activity (Chong et al., 2018). Even though *APBB2* (amyloid beta A4 precursor protein-binding, family B, member 2) primarily binds to *APP*, knocking out *APBB2* in mice causes glucose intolerance and  $\beta$  cell dysfunction (Ye et al., 2018). In transgenic mice, deleting *STAT3* in  $\beta$  cells and neurons impaired glucose metabolism (Cui et al., 2004). *STAT3* also regulates liver glucose homeostasis by modulating the expression of gluconeogenic genes (Inoue et al., 2004). A gene co-regulatory network analysis identified *RAB11FIP2* as a differentially expressed gene in axon regeneration, suggesting its possible role in atrophy (Su et al., 2018). Correspondingly, a growing body of evidence supports the gene-imaging interactions we found in longitudinal alterations in atrophy. *CAST* overexpression was shown to reduce amyloid burden due to its effect on *BACE1* processing of *APP* (Liang et al., 2010; Morales-Corraliza et al., 2012). *FHL2* prevents inflammatory angiogenesis and regulates the function of vascular smooth muscle cells, suggesting its role in blood flow (C. Y. Chen et al., 2020; Chu et al., 2008). *IGF2R* (insulin-like growth factor 2 receptor) interacts with insulin receptors for energy homeostasis, and the dysregulation of the gene is associated with type 2 diabetes (Chanprasertyothin et al., 2015). *RUVBL1* is an ATPase which modulates insulin signalling, and *RUVBL1* knock-out mice displayed impaired glucose metabolism (Mello et al., 2020).

## 2.5.2 Aging and Alzheimer's disease have both common and distinct mechanisms

In this study, we used a single gene expression template for all the subjects due to the unavailability of individual whole-brain gene expression. However, notice that even though this template has spatial but no temporal variation, for each gene, a model parameter controls its interaction (at the individual level) with each time-varying neuroimaging modality (i.e. the estimated transcriptomic-imaging parameters). At the individual level, the fitted gene-imaging parameters are assumed to reflect the gene-specific deformations required to fit the data. Consequently, these parameters represent quantitative measures of the individual dysregulation or deviation in gene expression patterns; and when analyzed across the entire population (e.g. via SVD analysis), the parameters can be used to detect cognitive/clinical related genetic associations.” Thus, under normal aging, the parameters obtained from the model optimization should be close to zero. Interestingly, it was observed that only ~70 parameters (out of over 35000 gene-imaging interaction parameters) were significantly different from zero across the healthy aging population. Conversely, ~1000 parameters significantly differed from zero across the diseased population. We attribute the greater number of significant parameters in AD to more genetic dysregulations and biological mechanism alterations in the disorder (Y. Iturria-Medina et al., 2020; Mostafavi et al., 2018).

The mechanisms of healthy aging and AD substantially overlap even though AD-related alterations are often accelerated, and the regions of alteration could be different (Toepper, 2017; Xia et al., 2018). Among the aging-associated genes, *CASP10*, *BIRC5*, and *PLSCR3* are involved in caspase-dependent apoptosis. Interestingly, apoptotic genes were also found in AD including *CASP3*, *CASP7*, *PLSCR1*, *CREB1*, *RELB*, *IGF2R*, *DFFB*. Sestrin (*SESNI*) is implicated in oxidative signalling, aging inhibition, and exercise mediation (Budanov et al., 2010; M. Kim et al., 2020; Yang et al., 2013). Correspondingly, some AD causal genes including *MEF2C*, *CBR1* and *NOSIP* are known for their roles in oxidative stress,

supporting the relevance of this pathway to both normal and pathological aging (Y. N. Kim et al., 2014; Rochette et al., 2013). Given that G-protein coupled receptors (GPCR) mediate the cellular response to most hormones/neurotransmitters (de Oliveira et al., 2019; Thathiah et al., 2011), it is unsurprising that GPCR-related genes converge on normal aging (*GNAI5*) and Alzheimer's disease (*GNAS*, *GNB5*). Having found some inflammation-associated genes in AD and the complement component *C5* in aging suggests that immune/inflammatory response change is part of both healthy aging and AD. Indeed, apart from the overlapping pathways, *LSM6* was the only gene common to both normal aging and AD. *LSM6* regulates gene expression and mRNA splicing, and a proteomic study linked its expression level to aging in human muscle cells (Ubaida-Mohien et al., 2019). Although altered mRNA splicing is associated with AD (Erik C. B. Johnson et al., 2018; Koch, 2018; Twine et al., 2011), a functional validation can further reveal the exact role of *LSM6* in the disease.

### **2.5.3 Towards a genetic approach to extending healthy aging and treating Alzheimer's disease**

The complexity of aging and the mixed aetiology of neurodegeneration necessitate an integrative multifactorial paradigm. In this study, we advanced the understanding of aging and AD pathology through the mechanistic modelling of how gene activity modulates relevant biological factors (e.g. tau, A $\beta$ , CBF, neuronal activity) to drive the cognitive alterations typically observed in the associated populations. The obtained results, in line with relevant molecular and imaging literature, highlight the strength of our approach by confirming previously identified aging- and AD-associated genes and uncovering new genes with relevant pathophysiological roles. In essence, this flexible formulation directly decodes the genetic mediators of spatiotemporal macroscopic brain alterations with aging and disease progression. Consequently, this work has important implications for the mechanistic understanding of aging

and AD pathogenesis and, importantly, for the implementation of a biologically defined patient stratification for personalized medical care.

Current approaches to AD treatment do not account for patient heterogeneity and such non-personalized methods may not only be ineffective but also cause undesired secondary effects in patients (Iturria-Medina et al., 2018). In a previous study, we used a similar imaging-based framework to show that some patients may need interventions targeting either tau, A $\beta$ , CBF or metabolism, while others can require a combinatorial therapy (e.g. concurrently targeting tau, A $\beta$ , and metabolic dysregulation) (Iturria-Medina et al., 2018). Based on this extended approach (GE-MCM), a gene therapy could replace the single and combinatorial treatment fingerprints described, by targeting highly influential genes modulating those factors in individuals. Many of the gene-imaging relationships found in our study have been previously reported *in vivo*, and the novel associations can be validated through experimental models. Understanding these relationships is crucial for effective drug development and administration. For instance, we found that *APBB2* is mediator of glucose metabolism. Thus, metabolic side effects may be considered when selecting *APBB2* as a therapeutic target of amyloid processing.

We have used inferred mRNA values for unobserved regions due to the unavailability of high-spatial resolution GE data. Nevertheless, the correlations between observed and predicted mRNA values are very high for majority of the genes (see Figure 2.S2), further supporting the feasibility of interpolating mRNA values based on spatial dependence (Gryglewski et al., 2018). It is however noteworthy that some genes with low correlation values might have low spatial dependence or error in the assay. There is inherent bias in the merged gene expression data from AHBA due to individual variability, and the AHBA subjects are not very representative of the typical age range in the ADNI cohort. Nevertheless, animal and human studies have reported large concordance between peripheral and brain gene expression, implying that blood gene expression may be used as a surrogate for gene expression in brain

tissue (Y. Iturria-Medina et al., 2020; Jasinska et al., 2009; Sullivan et al., 2006; Witt et al., 2013). Thus, our future work will focus on using personalized gene expression data from blood samples. The applicability and generalizability of the current formulation would also be tested in other neurological conditions (e.g. Parkinson’s disease and frontotemporal dementia).

## **2.6 Materials and methods**

### **2.6.1 Data description and processing**

#### ***Study participants***

This study involved 944 individuals with six multimodal brain imaging from ADNI (RRID:SCR\_003007) (<http://adni.loni.usc.edu/>; Figure 2.S1). First, for each imaging modality, a multivariate outlier identification was performed based on the Mahalanobis distance, with a significant squared distance ( $p < 0.05$ ) denoting an outlier (Iturria-Medina et al., 2016). From the 911 subjects that survived outlier detection, we chose 509 subjects having at least four imaging modalities (between amyloid PET, tau PET, glucose metabolism PET, resting-state fMRI, cerebral blood flow ASL, and structural MRI). Then, 460 subjects with at least three time points in any of the imaging modalities were selected for our analyses. Next, for each of these subjects ( $N = 460$ ), missing imaging modalities at each time point having actual individual data were automatically imputed using the trimmed scores regression with internal PCA (Folch-Fortuny et al., 2016). The accuracy of the imputation was validated with a leave-one-out cross-validation (e.g., tau imaging data can be significantly recovered for each subject with actual data,  $p < 10^{-6}$ ). Hence, all the 460 subjects used in subsequent analyses have completed all six neuroimaging modalities and an average of 4.7 ( $\pm 2.5$ ) longitudinal time points. Please see Figure 2.S1 for a detailed flowchart of subject selection and Table 2.S1 for demographic characteristics. Among the 460 participants, 151 were clinically identified as asymptomatic or HC, 161 with EMCI, 113 with LMCI, and 35 with probable AD.

### ***Whole-brain gene expression data and brain parcellation***

Microarray data was downloaded from the AHBA (RRID:SCR\_007416) website (<http://www.brain-map.org>) (Hawrylycz et al., 2012). The AHBA data consists of mRNA expression in 3702 tissue samples obtained from six neurotypical adult brains. The data were preprocessed by the Allen Institute to reduce the effects of bias due to batch effects. Description of the processing steps can be found in the technical white paper (Allen Human Brain Atlas, 2013). For each brain, there are 58,692 probes representing 20,267 unique genes. Transcriptome shows spatial dependence, with adjacent regions having similar expression pattern values (Gryglewski et al., 2018). Gaussian kernel regression affords a method of predicting gene expression values for unobserved regions based on the mRNA values of proximal regions. The regression is done as a weighted linear combination of unobserved mRNA, with the weight decreasing outward from proximal to distal regions. In order to select a representative probe for genes with multiple probes, Gaussian kernel regression was applied to predict the mRNA intensity in each of the 3702 samples in MNI space (Evans et al., 1994) using leave-one-out cross-validation. The probe with the highest prediction accuracy (among the multiple probes for a gene) was chosen as the representative probe for that gene. Next, because GE values were not available for all the grey matter voxels of the brain, Gaussian kernel regression was also used to predict the GE for the remaining MNI coordinates without mRNA expression intensity. Thus, the whole-brain GE data was obtained for the selected 20,267 probes/genes. It was infeasible to use these ~20,000 AHBA genes for modelling, we therefore selected 976 AHBA genes that can be found in the list of 978 landmark genes identified by Subramanian et al., 2017. These landmark genes are universally informative transcripts with the capacity to cover most of the information in the whole human transcriptome across a diversity of tissue types (Table 2.S2).

The brain was parcellated into 144 grey matter regions, and the average expression value of each gene was calculated for each region. The brain parcellation was derived from a combination of two atlases: 88 regions were identified through cytoarchitecture from Julich atlas (Palomero-Gallagher and Zilles, 2019) and 56 regions were derived from Brodmann atlas. Six regions were excluded due to zero or strong outlier PET imaging signals in their volumes. The remaining 138 regions were used for analyses (Table 2.S3).

### ***Cognitive and clinical evaluations***

The participants were characterized cognitively using MMSE, MEM, EF (Gibbons et al., 2012), and ADAS-Cognitive Subscales 11 and 13 (ADAS-11 and ADAS-13, respectively). They were also clinically diagnosed at baseline as HC, EMCI, LMCI, or probable AD patient.

### ***Multimodal imaging modalities***

#### **ASL MRI**

Resting arterial spin labeling (ASL) data were acquired using the Siemens product PICORE sequence (N = 213) with acquisition parameters TR/TE = 3400/12 ms, T11/TI2 = 700/1900 ms, FOV = 256 mm, 24 sequential 4-mm-thick slices with a 25% gap between the adjacent slices, partial Fourier factor = 6/8, bandwidth = 2368 Hz/pix, and imaging matrix size = 64 × 64. The data were processed in six steps as follows: (1) motion correction, (2) perfusion-weighted images (PWI) computation, (3) intensity scaling, (4) CBF image calculation, (5) spatial normalization to MNI space (Evans et al., 1994) using the registration parameters obtained for the structural T1 image with the nearest acquisition date, and (6) the mean CBF calculation for each of the considered brain regions. Details of the processing can be found at <http://www.adni.loni.usc.edu> under ‘UCSF ASL Perfusion Processing Methods’.

#### **Amyloid-β PET**

A 370-MBq bolus injection of AV-45 was administered to each subject and, after about 50 min, 20-min continuous brain PET imaging scans were acquired (N = 459). The images were reconstructed immediately after the scan and, when motion artifact was detected, another 20-min continuous scan was acquired. The acquired PET scans were then preprocessed using the following four main steps as described in Jagust et al., 2010: (1) dynamic co-registration to reduce motion artifacts, (2) across-time averaging, (3) re-sampling and reorientation of scans from native space to a standard voxel image grid space ('AC-PC' space), and (4) spatial filtering to convert the images to a uniform isotropic resolution of 8 mm FWHM. Finally, using the registration parameters obtained for the structural T1 image with the nearest acquisition date, all A $\beta$  scans were transformed to the MNI space (Evans et al., 1994). Using the cerebellum as an A $\beta$  non-specific binding reference, standardized uptake value ratio (SUVR) values were calculated for the 138 brain regions under consideration.

### **Resting-state fMRI**

Resting-state fMRI scans were acquired using an echo-planar pulse sequence on a 3.0T Philips MRI scanner (N = 148) with the following parameters: 140 time points, repetition time (TR) = 3000 ms, echo time (TE) = 30 ms, flip angle = 80°, number of slices = 48, slice thickness = 3.3 mm, spatial resolution = 3×3×3 mm<sup>3</sup>, and in-plane matrix size = 64 × 64. The scans were corrected for motion and slice timing. Then, they were spatially normalized to MNI space (Evans et al., 1994) using the registration parameters obtained for the structural T1 image with the nearest acquisition date. Signal filtering was performed to retain only low-frequency fluctuations (0.01–0.08 Hz) (Chao-Gan and Yu-Feng, 2010). Fractional amplitude of low-frequency fluctuation (fALFF) was calculated as a regional quantitative indicator of the brain's functional integrity. fALFF has been shown to be highly sensitive to disease progression (Iturria-Medina et al., 2016).

### **Fluorodeoxyglucose PET**

A 185-MBq (5 + 0.5 mCi) bolus injection of [18F]-FDG was administered to each subject and brain PET imaging data were obtained approximately 20 min after injection (N = 455). The images were attenuation-corrected and then preprocessed as follows Jagust et al., 2010: (1) dynamic co-registration of frames to reduce the effects of patient motion, (2) across-time averaging, (3) reorientation from native space to a standard voxel image grid ('AC-PC'), and (4) spatial filtering to convert the images to a uniform isotropic resolution of 8 mm FWHM. Next, using the registration parameters obtained for the structural T1-weighted image with nearest acquisition date, the FDG-PET images were normalized to the MNI space (Evans et al., 1994). The cerebellum was then used as a reference to calculate SUVR values for the 138 regions (Klein and Tourville, 2012).

### **Structural MRI**

Structural T1-weighted 3D images were obtained for all subjects (N = 460) as described in <http://adni.loni.usc.edu/methods/documents/mri-protocols/>. The images were corrected for intensity nonuniformity using the N3 algorithm (Sled et al., 1998). Next, they were segmented into grey matter (GM), white matter (WM), and cerebrospinal fluid (CSF) probabilistic maps, using SPM12 (<http://www.fil.ion.ucl.ac.uk/spm>). The grey matter segmentations were transformed into the MNI space (Evans et al., 1994) using DARTEL (Ashburner, 2007). To preserve the initial amount of tissue volume, each map was corrected for the effects of the spatial registration. Mean grey matter density and determinant of the Jacobian (DJ) (Ashburner, 2007) values were calculated for the 138 grey matter regions (Klein and Tourville, 2012). The grey matter density was used in this study as a measure of structural atrophy.

### **Tau PET**

A 370-MBq/kg bolus injection of tau-specific ligand 18F-AV-1451 ([F- 18] T807) was given to each subject, and 30-min (6 × 5 min frames) brain PET scans were acquired at 75 min after injection (N = 233). As previously described (Jagust et al., 2010), the images were

preprocessed as follows: (1) dynamic co-registration, (2) across-time averaging, (3) resampling and reorientation from native space to a standard voxel image grid space ('AC-PC' space), and (4) using ion parameters obtained for the structural T1 image with the nearest acquisition date, all tau images were normalized to the MNI space (Evans et al., 1994). The cerebellum was used as a reference to calculate SUVR values for the 138 grey matter regions.

### ***Anatomical connectivity estimation***

The connectivity matrix was constructed in DSI Studio (<http://dsi-studio.labsolver.org>) using a group average template from 1065 subjects (Yeh et al., 2018). A multi-shell high-angular-resolution diffusion scheme was used, and the b-values were 990, 1985, and 2980 s/mm<sup>2</sup>. The total number of sampling directions was 270. The in-plane resolution and slice thickness were 1.25 mm. The diffusion data were reconstructed in the MNI space using q-space diffeomorphic reconstruction to obtain the spin distribution function (Yeh and Tseng, 2011; Yeh et al., 2010). The sampling length and output resolution were set to 2.5 and 1 mm, respectively. The restricted diffusion was quantified using restricted diffusion imaging (Yeh et al., 2017) and a deterministic fibre tracking algorithm was used (Yeh et al., 2013). Using the brain atlas previously described under 'Methods' ('Whole-brain gene expression data and brain parcellation'), seeding was placed on the whole brain while setting the QA threshold to 0.15. The angular threshold was randomly varied from 15 to 90 degrees and the step size from 0.5 to 1.5 voxels. The fibre trajectories were smoothed by averaging the propagation direction with a percentage of the previous direction, which was randomly selected from 0 to 95%. Tracks with lengths shorter than 30 mm or longer than 300 mm were then discarded. A total of 100,000 tracts were calculated, and the connectivity matrix was obtained by using count of the connecting tracks.

## 2.6.2 Gene Expression Multifactorial Causal Model

In the basic MCM formulation (Iturria-Medina et al., 2017), the brain is considered as a dynamic multifactorial causal system, where: i) each variable represents a relevant macroscopic biological factor at a given brain region (e.g. tau and amyloid proteins, CBF, neuronal activity at rest, grey matter density), and ii) alterations in each biological factor are caused by direct factor-factor interactions, the intra-brain propagation of factor-specific alterations (e.g. tau and amyloid spreading), and external inputs (e.g. drugs). Here, we extend this approach to incorporate GE at the regional level. Specifically, we examine how macroscopic biological alterations at each brain region, and the associated macroscopic factor-factor interactions, are controlled by the regional genetic activity.

The GE-MCM is therefore defined by: (i) the influence of each gene on the local direct interactions among all the macroscopic factors, constrained within each brain region, and (ii) the potential spreading of macroscopic factor-specific alterations through anatomical and/or vascular networks. Mathematically, these processes can be described as:

$$\frac{dS_i^m}{dt} = \sum_{n=1}^{N_{factors}} \left( \alpha_o^{n \rightarrow m} + \sum_{k=1}^{N_{genes}} \alpha_k^{n \rightarrow m} G_i^k \right) S_i^n + \sum_{\substack{j=1 \\ j \neq i}}^{N_{rois}} C_{ji}^m (S_j^m - S_i^m) S^m \quad (1)$$

$N_{genes} = 976$  is the number of genes. Each gene was normalized by z-score across  $N_{rois} = 138$  brain grey matter regions of interest (a gene  $i$  is denoted as  $G_i$ ; for region names, see *Table 2.S3*).  $N_{factors} = 6$  is the number of biological factors measured at the same brain regions (i.e. A $\beta$  deposition, tau deposition, CBF, glucose metabolism, functional activity at rest, and grey matter density). Each node, corresponding to a given biological factor  $m$  and region  $i$ , is characterized by  $S_i^m \in \mathbb{R}$ .

In the equation,  $\frac{dS_i^m}{dt}$  is the local longitudinal alteration of a macroscopic factor  $m$  at region  $i$ , because of the foregoing multiscale interactions. The first term on the right models the local direct influences of multiple macroscopic biological factors on the given factor  $m$ .

The interaction parameters (  $\alpha_o^{n \rightarrow m}$ ,  $\alpha_k^{n \rightarrow m}$  ) and gene expression ( $G_i^k$ ) modulate the direct within-region impact of the factor  $n$  on  $m$ , including intra-factor effects, i.e. when  $n = m$ .

$\sum_{\substack{j=1 \\ j \neq i}}^{N_{rois}} C_{ji}^m (S_j^m - S_i^m) S^m$  reflects the resultant signal propagation of factor  $m$  from region  $i$  to

other brain regions through the physical network  $C_{ji}^m$ .

The GE-MCM model can advance our mechanistic understanding of the complex processes of aging and neurodegeneration. Its ability to map a healthy gene expression template to each subject allows us to model how the spatial distribution of transcriptome drives the multifactorial alteration observed in the brain. The interaction parameter  $\alpha_k^{n \rightarrow m}$  is an implicit quantitative measure of dysregulation or deviation of gene expression from normal patterns. By fitting the model at the individual level, it is possible to identify subject-specific genetic targets for personalized treatment of AD and enhancing healthy aging.

### 2.6.3 Model evaluation

The GE-MCM differential equation (1) was solved for each participant. For each subject  $j$  and biological factor  $m$ ,  $\frac{dS_i^m(j)}{dt}$  was calculated between each pair of consecutive time points, and the regional values obtained were concatenated into a subject- and factor-specific vector ( $\frac{dS^m(j)}{dt}$ ) with  $N_{rois} \cdot (N_{times} - 1)$  unique values. This concatenation allowed us to express the evaluation of the model parameters (  $\alpha_o^{n \rightarrow m}$ ,  $\alpha_k^{n \rightarrow m}$  ) as a regression problem (with  $\frac{dS^m(j)}{dt}$  as dependent variable). We applied a Bayesian sparse linear regression with horseshoe hierarchy to identify the distribution of the model parameters (Carvalho et al., 2010; Makalic et al., 2016). Due to high dimensionality of the data, a computationally efficient algorithm was used to sample the posterior Gaussian distribution of the regression coefficients (Bhattacharya et al., 2016), and the algorithm was implemented in MATLAB (Makalic et al., 2016). Through Markov chain Monte Carlo, we generated 500 samples of each regression coefficient after

discarding the first 1000 burn-in simulations. All 500 samples were averaged, and 5863 coefficients were obtained for every subject and biological factor. For subsequent analysis, we used 5856 coefficients (transcriptomic-imaging parameters) that corresponded to the measure of transcriptomic effect on the interaction of a macroscopic imaging-based factor with the other macroscopic factors, in driving a longitudinal biological factor alteration.

Next, we sought to identify the top genes mediating cognitive and behavioural changes in healthy aging and AD progression. First, we identified 113 clinically stable HC subjects who did not convert to MCI or AD stage within  $7.8 \pm 2.9$  years. In addition, we selected 129 diseased subjects diagnosed with AD at baseline or AD converters (i.e. HC and MCI subjects that advanced to AD within  $3.7 \pm 2.9$  years). For each independent subset of subjects (i.e. stable HC or diseased subjects), we combined the transcriptomic-imaging parameters across the six longitudinal biological factor alterations (see Figure 2.1D). We then evaluated the across-population stability of these model parameters via their 99% confidence intervals (99% CI). Next, rate of change of cognitive scores were calculated for each subject ( $7.2 \pm 2.6$  time points for HC and  $6.3 \pm 3.0$  time points for AD). We applied singular value decomposition (SVD) multivariate analysis to evaluate how the stable transcriptomic-imaging interactions mediate group-specific changes in cognitive/clinical scores (age-related slopes of MMSE, ADAS-11, ADAS-13, EM, and EF). For each group (i.e. HC or AD), SVD identified a few pairs of “principal components” that maximize the cross-correlation between the two sets of variables (Carbonell et al., 2020; Worsley et al., 2005). Then it mapped the gene-imaging parameters onto the obtained principal components (PC). This mapping provides a score (or contribution) of a gene-imaging parameter to a PC. Next, the significant PC were identified by running 10000 permutations. To identify the genes (gene-imaging parameters) with large and reliable contributions on the significant PC, we drew 10000 bootstrap samples and calculated the bootstrap ratio of the gene-imaging parameters. The bootstrap ratio is obtained by dividing the

gene-imaging saliences (contributions) by their respective bootstrap standard errors. It allowed us to assess the reliability of the genetic contributions (McIntosh et al., 2004). Hence, top aging- or AD-related causal genes were identified by selecting the parameters with bootstrap ratio above 2.58, which is approximately equivalent to a z-score for 99% CI if the bootstrap distribution is normal (Efron et al., 1986).

## **2.7 Data availability**

All data used in this study are publicly available at the Allen Human Brain Atlas website (Hawrylycz et al., 2012. *Nature*, 489:391-399; <http://human.brain-map.org/static/download>) and the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (Peterson et al., 2010. *Neurology*, 74(3): 201-209; <http://adni.loni.usc.edu/data-samples/access-data/>). While AHBA data do not require any registration for download, ADNI data can be accessed by creating an account and submitting an online application form. The application includes the investigator's institutional affiliation and the proposed uses of the ADNI data (scientific investigation, teaching, or planning clinical research studies). ADNI data may not be used for commercial products or redistributed in any way.

The following previously published data sets were used:

Hawrylycz et al. (2012) Allen Human Brain Atlas ID RRID:SCR\_007416. The Allen Human Brain Atlas. <http://human.brain-map.org/static/download>

Petersen et al. (2010) Alzheimer's Disease Neuroimaging Initiative ID RRID:SCR\_003007. The Alzheimer's Disease Neuroimaging Initiative. <http://adni.loni.usc.edu/data-samples/access-data/>

## 2.8 Supplementary Figures

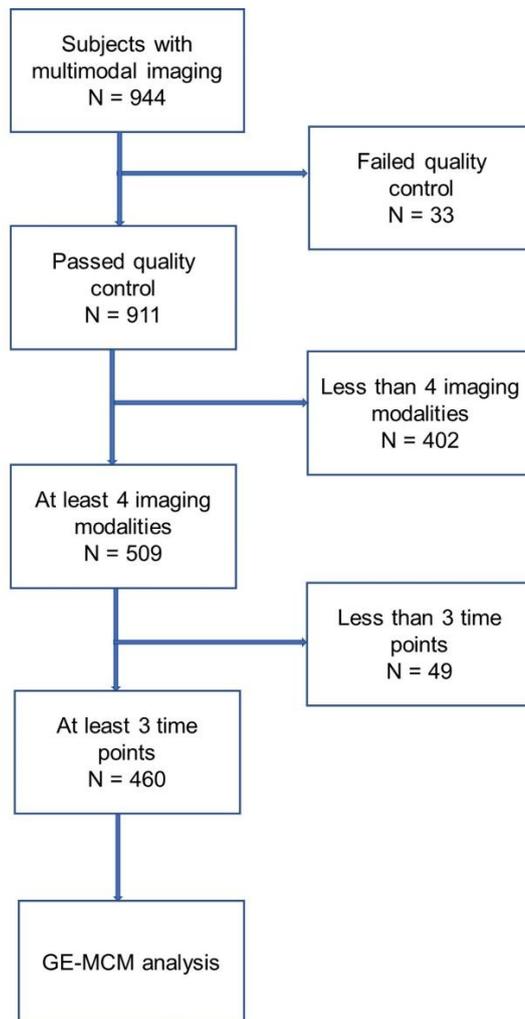


Figure 2.S1: Subject selection. A multivariate outlier identification was performed based on 944 subjects. From the 911 subjects that survived outlier detection, 509 subjects having at least four imaging modalities were chosen. Then, 460 subjects with at least three time points in any of the imaging modalities were selected. Next, for each of the 460 subjects, missing imaging modalities at each time point having actual individual data were automatically imputed using the trimmed scores regression with internal PCA.

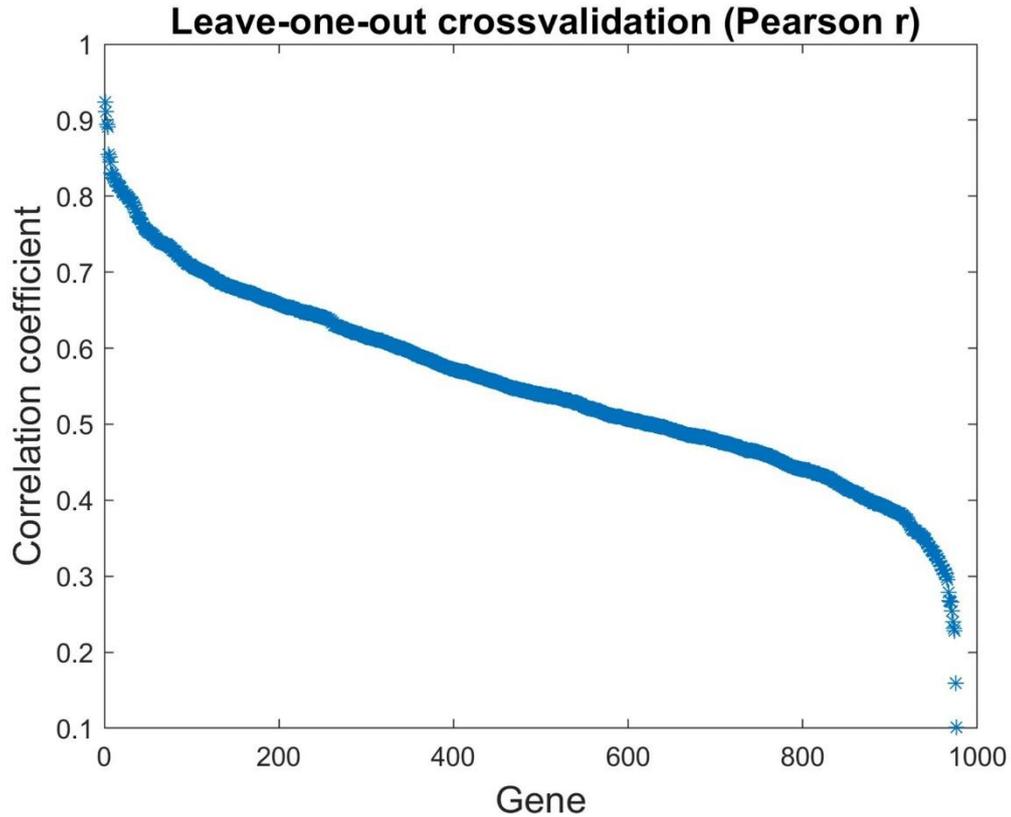


Figure 2.S2: Correlation of predicted mRNA expression with actual mRNA expression across 976 genes. Using the actual mRNA values of 3702 samples from AHBA, Gaussian kernel regression was used to reproduce the mRNA intensities through a leave-one-out cross-validation, and the Pearson correlation coefficient between the actual and reproduced values was calculated for each gene.

## 2.9 Supplementary Tables

Table 2.S1: Main demographic characteristics of the included ADNI subjects.

Variable	HC (N=151)	EMCI (N=161)	LMCI (N=113)	AD (N=35)	Stable HC (N=113)	AD + converters (N=129)
Female	76(50.3%)	68(42.2%)	51(45.1%)	16(45.7%)	59(52.2%)	58(45%)
Mean age (years)	74(5.5)	70.1(6.8)	71.7(7.1)	74.7(8.1)	73.7(5.6)	73.2(7.1)
Mean education (years)	16.5(2.7)	16.3(2.7)	16.2(2.9)	15.2(2.6)	16.8(2.5)	15.8(2.7)

Data are number (%) or mean (std).

Table 2.S2: List of 976 genes used in this study. Please see Excel file.

Table 2.S3: Brain regions used in this study.

Number	Julich Atlas	Number	Brodmann's Atlas
1	hOc1	44	Brodmann's area 1
2	hOc2	45	Brodmann's area 2
3	hOc4d	46	Brodmann's area 3
4	hOc3d	47	Brodmann's area 4
5	hOc3v	48	Brodmann's area 5
6	hOc4v	49	Brodmann's area 6
7	1	50	Brodmann's area 7
8	2	51	Brodmann's area 10
9	†3a	52	Brodmann's area 11
10	3b	53	Brodmann's area 17
11	FG1	54	†*Brodmann's area 18
12	FG2	55	Brodmann's area 19
13	Brodmann's area 37	56	Brodmann's area 24
14	Te1	57	Brodmann's area 25
15	Te2	58	Brodmann's area 26
16	Brodmann's area 20	59	Brodmann's area 27
17	Brodmann's area 21	60	Brodmann's area 29
18	Brodmann's area 22	61	†*Brodmann's area 30
19	Brodmann's area 36	62	Brodmann's area 32
20	Brodmann's area 38	63	Brodmann's area 34
21	5L	64	Brodmann's area 35
22	5M	65	Brodmann's area 39
23	PGa	66	Brodmann's area 40
24	PGp	67	†Brodmann's area 41
25	PFt	68	Brodmann's area 42
26	PFm	69	Brodmann's area 43
27	p24ab	70	Brodmann's area 44
28	p32	71	Brodmann's area 45
29	Brodmann's area 23	72	Brodmann's area 48
30	6		
31	4p		
32	Brodmann's area 8		
33	Brodmann's area 9		
34	Fp1		
35	Fp2		

Number	Julich Atlas	Number	Brodmann's Atlas
36	Fo1		
37	44		
38	45		
39	Brodmann's area 46		
40	Brodmann's area 47		
41	7A		
42	CA+dentate		
43	Brodmann's area 28		

Table 2.S4: Distribution of stable gene-imaging interaction parameters in healthy aging and AD progression (99% CI).

		Longitudinal biological factor alterations											
		CBF		A $\beta$		Functional activity		Glucose metabolism		Grey matter density		Tau	
		Healthy	AD	Healthy	AD	Healthy	AD	Healthy	AD	Healthy	AD	Healthy	AD
Gene-imaging interactions	CBF	0	17	5	34	3	28	2	19	0	13	3	31
	A $\beta$	0	74	4	40	4	20	0	19	1	15	2	17
	Functional activity	1	41	2	14	0	17	0	15	1	11	2	30
	Glucose metabolism	2	78	3	35	1	24	1	33	1	10	2	28
	Grey matter density	1	53	2	37	5	36	2	30	1	21	1	39
	Tau	1	10	4	17	3	15	0	9	0	10	8	53
	Total	5	273	20	177	16	140	5	125	4	80	18	198

Table 2.S5: Identified molecular pathways underlying AD progression.

Pathway	No of genes
CCKR signaling map	8
Inflammation mediated by chemokine and cytokine signaling pathway	6
Apoptosis signaling pathway	5
Gonadotropin-releasing hormone receptor pathway	5
Heterotrimeric G-protein signaling pathway-Gi alpha and Gs alpha mediated pathway	3
FAS signaling pathway	3
p38 MAPK pathway	3
Enkephalin release	3
Beta3 adrenergic receptor signaling pathway	2
Beta2 adrenergic receptor signaling pathway	2
Beta1 adrenergic receptor signaling pathway	2
5HT4 type receptor mediated signaling pathway	2

Angiogenesis	2
Alzheimer disease-presenilin pathway	2
Ubiquitin proteasome pathway	2
Wnt signaling pathway	2
N-acetylglucosamine metabolism	2
Cytoskeletal regulation by Rho GTPase	2
Histamine H2 receptor mediated signaling pathway	2
Cell cycle	2
B cell activation	2
Corticotropin releasing factor receptor signaling pathway	2
Axon guidance mediated by netrin	1
Axon guidance mediated by Slit/Robo	1
Metabotropic glutamate receptor group III pathway	1
JAK/STAT signaling pathway	1
Interleukin signaling pathway	1
Interferon-gamma signaling pathway	1
5HT2 type receptor mediated signaling pathway	1
Coenzyme A biosynthesis	1
5HT1 type receptor mediated signaling pathway	1
Insulin/IGF pathway-protein kinase B signaling cascade	1
Insulin/IGF pathway-mitogen activated protein kinase kinase/MAP kinase cascade	1
Huntington disease	1
Heterotrimeric G-protein signaling pathway-rod outer segment phototransduction	1
p53 pathway	1
p53 pathway feedback loops 2	1
Heterotrimeric G-protein signaling pathway-Gq alpha and Go alpha mediated pathway	1
p53 pathway by glucose deprivation	1
O-antigen biosynthesis	1
Xanthine and guanine salvage pathway	1
Transcription regulation by bZIP transcription factor	1
Thyrotropin-releasing hormone receptor signaling pathway	1
Toll receptor signaling pathway	1
Ras Pathway	1
Adenine and hypoxanthine salvage pathway	1
T cell activation	1
Oxytocin receptor mediated signaling pathway	1
Endothelin signaling pathway	1
EGF receptor signaling pathway	1
Parkinson disease	1
DNA replication	1
PI3 kinase pathway	1
Opioid proopiomelanocortin pathway	1
PDGF signaling pathway	1
Opioid prodynorphin pathway	1

Oxidative stress response	1
Opioid proenkephalin pathway	1
Cholesterol biosynthesis	1

## 2.10 Article information

### 2.10.1 Acknowledgement

This research was undertaken by funding from the Canada First Research Excellence Fund, awarded to McGill University for the Healthy Brains for Healthy Lives Initiative, the Fonds de la Recherche en Sante du Quebec (FRQS) Research Scholars Junior 1, the Canada Research Chair Tier-2 and the Weston Brain Institute Rapid Response programs 2018 and 2019 awards to YIM, and the Brain Canada Foundation and Health Canada support to the McConnell Brain Imaging Centre at the Montreal Neurological Institute. Dataset-1 collection and sharing was provided by Allen Institute for Brain Science (funded in part by the Department of Health and Human Services Health Resources and Services Administration awards 1C76HF15069-01-00 and 1C76HF19619-01-00). The Allen Human Brain Atlas is also supported by Paul G Allen, Mr. and Mrs. Richard Daly, Mr. and Mrs. Peter Eschenbach, Mr. William E Fay, Jr., Mr. Nathan Hansen, Mr. and Mrs. Richard Jernstedt, Bruce and Kathryn Johnson, Mr. and Mrs. Martin Koldyke, Michael and Suzanne Moskow, Bruce and Gwill Newman, Mr. and Mrs. Peter Pond, Mr. and Mrs. James N Bayer, Jr., Mr. and Mrs. Richard L Joutras, Mr. and Mrs. Robert Lorch, Lee and Jeanne Zehrer, and the board members of Brain Research Foundation.

The Dataset-2 collection and sharing for this project was funded by ADNI (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc;

Eisai, Inc; Elan Pharmaceuticals, Inc; Eli Lilly and Company; EuroImmune; F Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc; Fujirebio; GE Healthcare; IXICO Ltd; Janssen Alzheimer Immunotherapy Research and Development, LLC; Johnson and Johnson Pharmaceutical Research and Development, LLC; Lumosity; Lundbeck; Merck and Co, Inc; Meso Scale Diagnostics, LLC; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer, Inc; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (<http://www.fnih.org/>). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data is disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

### **2.10.2 Ethics**

Human subjects: This article does not contain any studies with human participants performed by any of the authors. The authors obtained approval from the ADNI Data Sharing and Publications Committee for data use and publication. As per ADNI protocols, the study was conducted according to Good Clinical Practice guidelines, the Declaration of Helsinki, US 21CFR Part 50 - Protection of Human Subjects, and Part 56 - Institutional Review Boards, and pursuant to state and federal HIPAA regulations ([adni.loni.usc.edu](http://adni.loni.usc.edu)). Study subjects and/or authorized representatives gave written informed consent at the time of enrollment for sample collection and completed questionnaires approved by each participating site Institutional Review Board (IRB).

## References

- Alkadhi, K., & Eriksen, J. (2011). The complex and multifactorial nature of Alzheimer's disease. *Current neuropharmacology*, 9(4), 586-586. doi:10.2174/157015911798376235
- Allen Human Brain Atlas. (2013). *Technical white paper: In situ hybridization in the Allen Human Brain Atlas*. [http://help.brain-map.org/download/attachments/2818165/ISH\\_WhitePaper.pdf?version=1&modificationDate=1382051868332&api=v2](http://help.brain-map.org/download/attachments/2818165/ISH_WhitePaper.pdf?version=1&modificationDate=1382051868332&api=v2)
- Anderson, C. M., Hu, J., Barnes, R. M., Heidt, A. B., Cornelissen, I., & Black, B. L. (2015). Myocyte enhancer factor 2C function in skeletal muscle is required for normal growth and glucose metabolism in mice. *Skeletal muscle*, 5, 7-7. doi:10.1186/s13395-015-0031-0
- Ashburner, J. (2007). A fast diffeomorphic image registration algorithm. *Neuroimage*, 38(1), 95-113. doi:10.1016/j.neuroimage.2007.07.007
- Austin, S. A., Santhanam, A. V., & Katusic, Z. S. (2010). Endothelial nitric oxide modulates expression and processing of amyloid precursor protein. *Circulation research*, 107(12), 1498-1502. doi:10.1161/CIRCRESAHA.110.233080
- Bezzi, P., Domercq, M., Brambilla, L., Galli, R., Schols, D., De Clercq, E., . . . Volterra, A. (2001). CXCR4-activated astrocyte glutamate release via TNF $\alpha$ : amplification by microglia triggers neurotoxicity. *Nature Neuroscience*, 4(7), 702-710. doi:10.1038/89490
- Bhattacharya, A., Chakraborty, A., & Mallick, B. K. (2016). Fast sampling with Gaussian scale-mixture priors in high-dimensional regression. *Biometrika*, 103(4), 985-991. doi:10.1093/biomet/asw042
- Budanov, A. V., Lee, J. H., & Karin, M. (2010). Stressin' Sestrins take an aging fight. *EMBO Mol Med*, 2(10), 388-400. doi:10.1002/emmm.201000097
- Carbonell, F., Zijdenbos, A. P., & Bedell, B. J. (2020). Spatially Distributed Amyloid- $\beta$  Reduces Glucose Metabolism in Mild Cognitive Impairment. *J Alzheimers Dis*, 73(2), 543-557. doi:10.3233/jad-190560
- Carvalho, C. M., Polson, N. G., & Scott, J. G. (2010). The horseshoe estimator for sparse signals. *Biometrika*, 97(2), 465-480.
- Chanprasertyothin, S., Jongjaroenprasert, W., & Ongphiphadhanakul, B. (2015). The association of soluble IGF2R and IGF2R gene polymorphism with type 2 diabetes. *Journal of diabetes research*, 2015, 216383-216383. doi:10.1155/2015/216383
- Chao-Gan, Y., & Yu-Feng, Z. (2010). DPARSF: A MATLAB Toolbox for "Pipeline" Data Analysis of Resting-State fMRI. *Front Syst Neurosci*, 4, 13. doi:10.3389/fnsys.2010.00013

- Chen, C. Y., Tsai, H. Y., Tsai, S. H., Chu, P. H., Huang, P. H., Chen, J. W., & Lin, S. J. (2020). Deletion of the FHL2 gene attenuates intima-media thickening in a partially ligated carotid artery ligated mouse model. *J Cell Mol Med*, 24(1), 160-173. doi:10.1111/jcmm.14687
- Chen, S.-D., Yang, J.-L., Lin, T.-K., & Yang, D.-I. (2019). Emerging Roles of Sestrins in Neurodegenerative Diseases: Counteracting Oxidative Stress and Beyond. *Journal of clinical medicine*, 8(7), 1001. doi:10.3390/jcm8071001
- Chong, Y., Saviuk, N., Pie, B., Basisty, N., Quinn, R. K., Schilling, B., . . . Haghghi, A. P. (2018). Removing 4E-BP Enables Synapses to Refine without Postsynaptic Activity. *Cell Rep*, 23(1), 11-22. doi:10.1016/j.celrep.2018.03.040
- Chu, P. H., Yeh, L. K., Lin, H. C., Jung, S. M., Ma, D. H., Wang, I. J., . . . Chen, J. (2008). Deletion of the FHL2 gene attenuating neovascularization after corneal injury. *Invest Ophthalmol Vis Sci*, 49(12), 5314-5318. doi:10.1167/iovs.08-2209
- Cui, Y., Huang, L., Eleftheriou, F., Yang, G., Shelton, J. M., Giles, J. E., . . . Li, C. (2004). Essential role of STAT3 in body weight and glucose homeostasis. *Molecular and cellular biology*, 24(1), 258-269. doi:10.1128/mcb.24.1.258-269.2004
- D'Amelio, M., Cavallucci, V., & Cecconi, F. (2010). Neuronal caspase-3 signaling: not only cell death. *Cell Death & Differentiation*, 17(7), 1104-1114. doi:10.1038/cdd.2009.180
- de Oliveira, P. G., Ramos, M. L. S., Amaro, A. J., Dias, R. A., & Vieira, S. I. (2019). G(i/o)-Protein Coupled Receptors in the Aging Brain. *Frontiers in aging neuroscience*, 11, 89-89. doi:10.3389/fnagi.2019.00089
- DeTure, M. A., & Dickson, D. W. (2019). The neuropathological diagnosis of Alzheimer's disease. *Molecular neurodegeneration*, 14(1), 32. doi:10.1186/s13024-019-0333-5
- Dillman, A. A., Majounie, E., Ding, J., Gibbs, J. R., Hernandez, D., Arepalli, S., . . . Cookson, M. R. (2017). Transcriptomic profiling of the human brain reveals that altered synaptic gene expression is associated with chronological aging. *Sci Rep*, 7(1), 16890. doi:10.1038/s41598-017-17322-0
- Dreyer, J., Schleicher, M., Tappe, A., Schilling, K., Kuner, T., Kusumawidijaja, G., . . . Kuner, R. (2004). Nitric oxide synthase (NOS)-interacting protein interacts with neuronal NOS and regulates its distribution and activity. *J Neurosci*, 24(46), 10454-10465. doi:10.1523/jneurosci.2265-04.2004
- Dukart, J., Kherif, F., Mueller, K., Adaszewski, S., Schroeter, M. L., Frackowiak, R. S. J., . . . Alzheimer's Disease Neuroimaging, I. (2013). Generative FDG-PET and MRI model of aging and disease progression in Alzheimer's disease. *PLoS computational biology*, 9(4), e1002987-e1002987. doi:10.1371/journal.pcbi.1002987
- Efron, B., & Tibshirani, R. (1986). Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy. *Statist. Sci.*, 1(1), 54-75. doi:10.1214/ss/1177013815
- Evans, A. C., Kamber, M., Collins, D. L., & MacDonald, D. (1994). An MRI-Based Probabilistic Atlas of Neuroanatomy. In S. D. Shorvon, D. R. Fish, F. Andermann, G.

- M. Bydder, & H. Stefan (Eds.), *Magnetic Resonance Scanning and Epilepsy* (pp. 263-274). Boston, MA: Springer US.
- Folch-Fortuny, A., Arteaga, F., & Ferrer, A. (2016). Missing Data Imputation Toolbox for MATLAB. *Chemometrics and Intelligent Laboratory Systems*, *154*, 93-100. doi:<https://doi.org/10.1016/j.chemolab.2016.03.019>
- Freeze, B., Acosta, D., Pandya, S., Zhao, Y., & Raj, A. (2018). Regional expression of genes mediating trans-synaptic alpha-synuclein transfer predicts regional atrophy in Parkinson disease. *Neuroimage Clin*, *18*, 456-466. doi:10.1016/j.nicl.2018.01.009
- Freeze, B., Pandya, S., Zeighami, Y., & Raj, A. (2019). Regional transcriptional architecture of Parkinson's disease pathogenesis and network spread. *Brain*, *142*(10), 3072-3085. doi:10.1093/brain/awz223
- Ge, J., Chen, Q., Liu, B., Wang, L., Zhang, S., & Ji, B. (2017). Knockdown of Rab21 inhibits proliferation and induces apoptosis in human glioma cells. *Cell Mol Biol Lett*, *22*, 30. doi:10.1186/s11658-017-0062-0
- Gibbons, L. E., Carle, A. C., Mackin, R. S., Harvey, D., Mukherjee, S., Insel, P., . . . Alzheimer's Disease Neuroimaging, I. (2012). A composite score for executive functioning, validated in Alzheimer's Disease Neuroimaging Initiative (ADNI) participants with baseline mild cognitive impairment. *Brain Imaging Behav*, *6*(4), 517-527. doi:10.1007/s11682-012-9176-1
- Giustiniani, J., Guillemeau, K., Dounane, O., Sardin, E., Huvent, I., Schmitt, A., . . . Chambraud, B. (2015). The FK506-binding protein FKBP52 in vitro induces aggregation of truncated Tau forms with prion-like behavior. *Faseb j*, *29*(8), 3171-3181. doi:10.1096/fj.14-268243
- Giustiniani, J., Sineus, M., Sardin, E., Dounane, O., Panchal, M., Sazdovitch, V., . . . Baulieu, E. E. (2012). Decrease of the immunophilin FKBP52 accumulation in human brains of Alzheimer's disease and FTDP-17. *J Alzheimers Dis*, *29*(2), 471-483. doi:10.3233/jad-2011-111895
- Gryglewski, G., Seiger, R., James, G. M., Godbersen, G. M., Komorowski, A., Unterholzner, J., . . . Lanzenberger, R. (2018). Spatial analysis and high resolution mapping of the human whole-brain transcriptome for integrative analysis in neuroimaging. *Neuroimage*, *176*, 259-267. doi:10.1016/j.neuroimage.2018.04.068
- Guerreiro, R., & Bras, J. (2015). The age factor in Alzheimer's disease. *Genome medicine*, *7*, 106-106. doi:10.1186/s13073-015-0232-5
- Hawrylycz, M. J., Lein, E. S., Guillozet-Bongaarts, A. L., Shen, E. H., Ng, L., Miller, J. A., . . . Jones, A. R. (2012). An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature*, *489*(7416), 391-399. doi:10.1038/nature11405
- Homolak, J., Mudrovčić, M., Vukić, B., & Toljan, K. (2018). Circadian Rhythm and Alzheimer's Disease. *Medical sciences (Basel, Switzerland)*, *6*(3), 52. doi:10.3390/medsci6030052

- Iaccarino, L., Tammewar, G., Ayakta, N., Baker, S. L., Bejanin, A., Boxer, A. L., . . . Rabinovici, G. D. (2017). Local and distant relationships between amyloid, tau and neurodegeneration in Alzheimer's Disease. *NeuroImage. Clinical*, *17*, 452-464. doi:10.1016/j.nicl.2017.09.016
- Inoue, H., Ogawa, W., Ozaki, M., Haga, S., Matsumoto, M., Furukawa, K., . . . Kasuga, M. (2004). Role of STAT-3 in regulation of hepatic gluconeogenic genes and carbohydrate metabolism in vivo. *Nature Medicine*, *10*(2), 168-174. doi:10.1038/nm980
- Ittner, L. M., & Götz, J. (2011). Amyloid- $\beta$  and tau--a toxic pas de deux in Alzheimer's disease. *Nat Rev Neurosci*, *12*(2), 65-72. doi:10.1038/nrn2967
- Iturria-Medina, Y., Carbonell, F., Assadi, A., Adewale, Q., Khan, A. F., Baumeister, R., & Sanchez-Rodriguez, L. M. (2020). NeuroPM toolbox: integrating Molecular, Neuroimaging and Clinical data for Characterizing Neuropathological Progression and Individual Therapeutic Needs. *medRxiv*, 2020.2009.2024.20200964. doi:10.1101/2020.09.24.20200964
- Iturria-Medina, Y., Carbonell, F. M., & Evans, A. C. (2018). Multimodal imaging-based therapeutic fingerprints for optimizing personalized interventions: Application to neurodegeneration. *Neuroimage*, *179*, 40-50. doi:10.1016/j.neuroimage.2018.06.028
- Iturria-Medina, Y., Carbonell, F. M., Sotero, R. C., Chouinard-Decorte, F., Evans, A. C., & Alzheimer's Disease Neuroimaging, I. (2017). Multifactorial causal model of brain (dis)organization and therapeutic intervention: Application to Alzheimer's disease. *Neuroimage*, *152*, 60-77. doi:10.1016/j.neuroimage.2017.02.058
- Iturria-Medina, Y., Khan, A. F., Adewale, Q., Shirazi, A. H., & Alzheimer's Disease Neuroimaging, I. (2020). Blood and brain gene expression trajectories mirror neuropathology and clinical deterioration in neurodegeneration. *Brain*, *143*(2), 661-673. doi:10.1093/brain/awz400
- Iturria-Medina, Y., Sotero, R. C., Toussaint, P. J., Mateos-Perez, J. M., Evans, A. C., & Alzheimer's Disease Neuroimaging, I. (2016). Early role of vascular dysregulation on late-onset Alzheimer's disease based on multifactorial data-driven analysis. *Nat Commun*, *7*, 11934. doi:10.1038/ncomms11934
- Jack, C. R., Jr., Wiste, H. J., Schwarz, C. G., Lowe, V. J., Senjem, M. L., Vemuri, P., . . . Petersen, R. C. (2018). Longitudinal tau PET in ageing and Alzheimer's disease. *Brain*, *141*(5), 1517-1528. doi:10.1093/brain/awy059
- Jagust, W. J., Bandy, D., Chen, K., Foster, N. L., Landau, S. M., Mathis, C. A., . . . Alzheimer's Disease Neuroimaging, I. (2010). The Alzheimer's Disease Neuroimaging Initiative positron emission tomography core. *Alzheimers Dement*, *6*(3), 221-229. doi:10.1016/j.jalz.2010.03.003
- Jasinska, A. J., Service, S., Choi, O. W., DeYoung, J., Grujic, O., Kong, S. Y., . . . Freimer, N. B. (2009). Identification of brain transcriptional variation reproduced in peripheral blood: an approach for mapping brain expression traits. *Hum Mol Genet*, *18*(22), 4415-4427. doi:10.1093/hmg/ddp397

- Johnson, E. C. B., Dammer, E. B., Duong, D. M., Ping, L., Zhou, M., Yin, L., . . . Seyfried, N. T. (2020). Large-scale proteomic analysis of Alzheimer's disease brain and cerebrospinal fluid reveals early changes in energy metabolism associated with microglia and astrocyte activation. *Nat Med*, *26*(5), 769-780. doi:10.1038/s41591-020-0815-6
- Johnson, E. C. B., Dammer, E. B., Duong, D. M., Yin, L., Thambisetty, M., Troncoso, J. C., . . . Seyfried, N. T. (2018). Deep proteomic network analysis of Alzheimer's disease brain reveals alterations in RNA binding proteins and RNA splicing associated with disease. *Molecular neurodegeneration*, *13*(1), 52-52. doi:10.1186/s13024-018-0282-4
- Khan, M. Z., Brandimarti, R., Shimizu, S., Nicolai, J., Crowe, E., & Meucci, O. (2008). The chemokine CXCL12 promotes survival of postmitotic neurons by regulating Rb protein. *Cell death and differentiation*, *15*(10), 1663-1672. doi:10.1038/cdd.2008.95
- Kim, M., Sujkowski, A., Namkoong, S., Gu, B., Cobb, T., Kim, B., . . . Lee, J. H. (2020). Sestrins are evolutionarily conserved mediators of exercise benefits. *Nature Communications*, *11*(1), 190. doi:10.1038/s41467-019-13442-5
- Kim, Y. N., Jung, H. Y., Eum, W. S., Kim, D. W., Shin, M. J., Ahn, E. H., . . . Choi, S. Y. (2014). Neuroprotective effects of PEP-1-carbonyl reductase 1 against oxidative-stress-induced ischemic neuronal cell damage. *Free Radic Biol Med*, *69*, 181-196. doi:10.1016/j.freeradbiomed.2014.01.006
- Klein, A., & Tourville, J. (2012). 101 labeled brain images and a consistent human cortical labeling protocol. *Front Neurosci*, *6*, 171. doi:10.3389/fnins.2012.00171
- Koch, L. (2018). Altered splicing in Alzheimer transcriptomes. *Nature Reviews Genetics*, *19*(12), 738-739. doi:10.1038/s41576-018-0064-4
- Kowald, A., & Kirkwood, T. B. (1996). A network theory of ageing: the interactions of defective mitochondria, aberrant proteins, free radicals and scavengers in the ageing process. *Mutat Res*, *316*(5-6), 209-236. doi:10.1016/s0921-8734(96)90005-3
- Kuintzle, R. C., Chow, E. S., Westby, T. N., Gvakharia, B. O., Giebultowicz, J. M., & Hendrix, D. A. (2017). Circadian deep sequencing reveals stress-response genes that adopt robust rhythmic expression during aging. *Nature Communications*, *8*, 14529-14529. doi:10.1038/ncomms14529
- Kurien, P., Hsu, P. K., Leon, J., Wu, D., McMahon, T., Shi, G., . . . Ptáček, L. J. (2019). TIMELESS mutation alters phase responsiveness and causes advanced sleep phase. *Proc Natl Acad Sci U S A*, *116*(24), 12045-12053. doi:10.1073/pnas.1819110116
- Lee, T. I., & Young, R. A. (2013). Transcriptional regulation and its misregulation in disease. *Cell*, *152*(6), 1237-1251. doi:10.1016/j.cell.2013.02.014
- Li, L., Xiao, L., Hou, Y., He, Q., Zhu, J., Li, Y., . . . Zhao, Y. (2016). Sestrin2 Silencing Exacerbates Cerebral Ischemia/Reperfusion Injury by Decreasing Mitochondrial Biogenesis through the AMPK/PGC-1 $\alpha$  Pathway in Rats. *Sci Rep*, *6*, 30272. doi:10.1038/srep30272

- Liang, B., Duan, B. Y., Zhou, X. P., Gong, J. X., & Luo, Z. G. (2010). Calpain activation promotes BACE1 expression, amyloid precursor protein processing, and amyloid plaque formation in a transgenic mouse model of Alzheimer disease. *J Biol Chem*, *285*(36), 27737-27744. doi:10.1074/jbc.M110.117960
- Makalic, E., & Schmidt, D. F. (2016). A simple sampler for the horseshoe estimator. *IEEE Signal Processing Letters*, *23*(1), 179-182. doi:10.1109/LSP.2015.2503725
- McIntosh, A. R., & Lobaugh, N. J. (2004). Partial least squares analysis of neuroimaging data: applications and advances. *Neuroimage*, *23 Suppl 1*, S250-263. doi:10.1016/j.neuroimage.2004.07.020
- Mei, Z., Yan, P., Tan, X., Zheng, S., & Situ, B. (2015). Transcriptional Regulation of BACE1 by NFAT3 Leads to Enhanced Amyloidogenic Processing. *Neurochemical Research*, *40*(4), 829-836. doi:10.1007/s11064-015-1533-1
- Mello, T., Materozzi, M., Zanieri, F., Simeone, I., Ceni, E., Bereshchenko, O., . . . Galli, A. (2020). Liver haploinsufficiency of RuvBL1 causes hepatic insulin resistance and enhances hepatocellular carcinoma progression. *Int J Cancer*, *146*(12), 3410-3422. doi:10.1002/ijc.32787
- Mi, H., Muruganujan, A., Casagrande, J. T., & Thomas, P. D. (2013). Large-scale gene function analysis with the PANTHER classification system. *Nature protocols*, *8*(8), 1551-1566. doi:10.1038/nprot.2013.092
- Morales-Corraliza, J., Berger, J. D., Mazzella, M. J., Veeranna, Neubert, T. A., Ghiso, J., . . . Mathews, P. M. (2012). Calpastatin modulates APP processing in the brains of  $\beta$ -amyloid depositing but not wild-type mice. *Neurobiology of aging*, *33*(6), 1125.e1129-1125.e1121.1125E1118. doi:10.1016/j.neurobiolaging.2011.11.023
- Mostafavi, S., Gaiteri, C., Sullivan, S. E., White, C. C., Tasaki, S., Xu, J., . . . De Jager, P. L. (2018). A molecular network of the aging human brain provides insights into the pathology and cognitive decline of Alzheimer's disease. *Nature Neuroscience*, *21*(6), 811-819. doi:10.1038/s41593-018-0154-9
- Niccoli, T., & Partridge, L. (2012). Ageing as a risk factor for disease. *Current biology : CB*, *22*(17), 741-752. doi:10.1016/j.cub.2012.07.024
- Palomero-Gallagher, N., & Zilles, K. (2019). Cortical layers: Cyto-, myelo-, receptor- and synaptic architecture in human cortical areas. *Neuroimage*, *197*, 716-741. doi:10.1016/j.neuroimage.2017.08.035
- Pascoal, T. A., Mathotaarachchi, S., Mohades, S., Benedet, A. L., Chung, C. O., Shin, M., . . . Rosa-Neto, P. (2017). Amyloid- $\beta$  and hyperphosphorylated tau synergy drives metabolic decline in preclinical Alzheimer's disease. *Molecular Psychiatry*, *22*(2), 306-311. doi:10.1038/mp.2016.37
- Pickett, E. K., Herrmann, A. G., McQueen, J., Abt, K., Dando, O., Tulloch, J., . . . Spires-Jones, T. L. (2019). Amyloid Beta and Tau Cooperate to Cause Reversible Behavioral and Transcriptional Deficits in a Model of Alzheimer's Disease. *Cell Rep*, *29*(11), 3592-3604.e3595. doi:10.1016/j.celrep.2019.11.044

- Prudente, S., Sesti, G., Pandolfi, A., Andreozzi, F., Consoli, A., & Trischitta, V. (2012). The mammalian tribbles homolog TRIB3, glucose homeostasis, and cardiovascular diseases. *Endocrine reviews*, *33*(4), 526-546. doi:10.1210/er.2011-1042
- Qosa, H., Abuasal, B. S., Romero, I. A., Weksler, B., Couraud, P. O., Keller, J. N., & Kaddoumi, A. (2014). Differences in amyloid- $\beta$  clearance across mouse and human blood-brain barrier models: kinetic analysis and mechanistic modeling. *Neuropharmacology*, *79*, 668-678. doi:10.1016/j.neuropharm.2014.01.023
- Rochette, L., Lorin, J., Zeller, M., Guillard, J. C., Lorgis, L., Cottin, Y., & Vergely, C. (2013). Nitric oxide synthase inhibition and oxidative stress in cardiovascular diseases: possible therapeutic targets? *Pharmacol Ther*, *140*(3), 239-257. doi:10.1016/j.pharmthera.2013.07.004
- Rodrigue, K. M., Kennedy, K. M., Devous, M. D., Sr., Rieck, J. R., Hebrank, A. C., Diaz-Arrastia, R., . . . Park, D. C. (2012).  $\beta$ -Amyloid burden in healthy aging: regional distribution and cognitive consequences. *Neurology*, *78*(6), 387-395. doi:10.1212/WNL.0b013e318245d295
- Rodríguez-Rodero, S., Fernández-Morera, J. L., Menéndez-Torre, E., Calvanese, V., Fernández, A. F., & Fraga, M. F. (2011). Aging genetics and aging. *Aging and disease*, *2*(3), 186-195. Retrieved from <https://pubmed.ncbi.nlm.nih.gov/22396873>
- <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3295054/>
- Sahu, S. K., Gummadi, S. N., Manoj, N., & Aradhyam, G. K. (2007). Phospholipid scramblases: An overview. *Archives of Biochemistry and Biophysics*, *462*(1), 103-114. doi:<https://doi.org/10.1016/j.abb.2007.04.002>
- Sao, T., Yoshino, Y., Yamazaki, K., Ozaki, Y., Mori, Y., Ochi, S., . . . Ueno, S. I. (2018). MEF2C mRNA expression and cognitive function in Japanese patients with Alzheimer's disease. *Psychiatry Clin Neurosci*, *72*(3), 160-167. doi:10.1111/pcn.12618
- Shi, X., Xu, L., Malagult, J., Tang, J., Yan, M., & Zhang, J. H. (2017). Sestrins: A New Kid for Stroke Treatment? *Curr Drug Deliv*, *14*(6), 797-806. doi:10.2174/156720181466616111125249
- Singh, P. P., Demmitt, B. A., Nath, R. D., & Brunet, A. (2019). The Genetics of Aging: A Vertebrate Perspective. *Cell*, *177*(1), 200-220. doi:10.1016/j.cell.2019.02.038
- Sled, J. G., Zijdenbos, A. P., & Evans, A. C. (1998). A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans Med Imaging*, *17*(1), 87-97. doi:10.1109/42.668698
- Su, L.-N., Song, X.-Q., Xue, Z.-X., Zheng, C.-Q., Yin, H.-F., & Wei, H.-P. (2018). Network analysis of microRNAs, transcription factors, and target genes involved in axon regeneration. *Journal of Zhejiang University. Science. B*, *19*(4), 293-304. doi:10.1631/jzus.B1700179
- Subramanian, A., Narayan, R., Corsello, S. M., Peck, D. D., Natoli, T. E., Lu, X., . . . Golub, T. R. (2017). A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. *Cell*, *171*(6), 1437-1452.e1417. doi:10.1016/j.cell.2017.10.049

- Sullivan, P. F., Fan, C., & Perou, C. M. (2006). Evaluating the comparability of gene expression in blood and brain. *Am J Med Genet B Neuropsychiatr Genet*, *141b*(3), 261-268. doi:10.1002/ajmg.b.30272
- Tanaka, T., Biancotto, A., Moaddel, R., Moore, A. Z., Gonzalez-Freire, M., Aon, M. A., . . . Ferrucci, L. (2018). Plasma proteomic signature of age in healthy humans. *Aging cell*, *17*(5), e12799-e12799. doi:10.1111/accel.12799
- Taneera, J., Dhaiban, S., Mohammed, A. K., Mukhopadhyay, D., Aljaibeji, H., Sulaiman, N., . . . Salehi, A. (2019). GNAS gene is an important regulator of insulin secretory capacity in pancreatic  $\beta$ -cells. *Gene*, *715*, 144028. doi:<https://doi.org/10.1016/j.gene.2019.144028>
- Thathiah, A., & De Strooper, B. (2011). The role of G protein-coupled receptors in the pathology of Alzheimer's disease. *Nature Reviews Neuroscience*, *12*(2), 73-87. doi:10.1038/nrn2977
- Toepper, M. (2017). Dissociating Normal Aging from Alzheimer's Disease: A View from Cognitive Neuroscience. *Journal of Alzheimer's disease : JAD*, *57*(2), 331-352. doi:10.3233/JAD-161099
- Twine, N. A., Janitz, K., Wilkins, M. R., & Janitz, M. (2011). Whole transcriptome sequencing reveals gene expression and splicing differences in brain regions affected by Alzheimer's disease. *PLoS One*, *6*(1), e16266. doi:10.1371/journal.pone.0016266
- Ubaida-Mohien, C., Lyashkov, A., Gonzalez-Freire, M., Tharakan, R., Shardell, M., Moaddel, R., . . . Ferrucci, L. (2019). Discovery proteomics in aging human skeletal muscle finds change in spliceosome, immunity, proteostasis and mitochondria. *Elife*, *8*. doi:10.7554/eLife.49874
- Wang, C., Yue, H., Hu, Z., Shen, Y., Ma, J., Li, J., . . . Gu, Y. (2020). Microglia mediate forgetting via complement-dependent synaptic elimination. *Science*, *367*(6478), 688-694. doi:10.1126/science.aaz2288
- Witt, S. H., Sommer, W. H., Hansson, A. C., Sticht, C., Rietschel, M., & Witt, C. C. (2013). Comparison of gene expression profiles in the blood, hippocampus and prefrontal cortex of rats. *In Silico Pharmacology*, *1*(1), 15. doi:10.1186/2193-9616-1-15
- Worsley, K. J., Chen, J.-I., Lerch, J., & Evans, A. C. (2005). Comparing Functional Connectivity via Thresholding Correlations and Singular Value Decomposition. *Philosophical Transactions: Biological Sciences*, *360*(1457), 913-920.
- Xia, X., Jiang, Q., McDermott, J., & Han, J.-D. J. (2018). Aging and Alzheimer's disease: Comparison and associations from molecular to system level. *Aging cell*, *17*(5), e12802-e12802. doi:10.1111/accel.12802
- Xu, Y., Liu, C., Chen, S., Ye, Y., Guo, M., Ren, Q., . . . Chen, L. (2014). Activation of AMPK and inactivation of Akt result in suppression of mTOR-mediated S6K1 and 4E-BP1 pathways leading to neuronal cell death in in vitro models of Parkinson's disease. *Cellular Signalling*, *26*(8), 1680-1689. doi:<https://doi.org/10.1016/j.cellsig.2014.04.009>

- Yang, Y. L., Loh, K. S., Liou, B. Y., Chu, I. H., Kuo, C. J., Chen, H. D., & Chen, C. S. (2013). SESN-1 is a positive regulator of lifespan in *Caenorhabditis elegans*. *Exp Gerontol*, 48(3), 371-379. doi:10.1016/j.exger.2012.12.011
- Ye, R., Gordillo, R., Shao, M., Onodera, T., Chen, Z., Chen, S., . . . Scherer, P. E. (2018). Intracellular lipid metabolism impairs  $\beta$  cell compensation during diet-induced obesity. *The Journal of clinical investigation*, 128(3), 1178-1189. doi:10.1172/JCI97702
- Yeh, F. C., Liu, L., Hitchens, T. K., & Wu, Y. L. (2017). Mapping immune cell infiltration using restricted diffusion MRI. *Magn Reson Med*, 77(2), 603-612. doi:10.1002/mrm.26143
- Yeh, F. C., Panesar, S., Fernandes, D., Meola, A., Yoshino, M., Fernandez-Miranda, J. C., . . . Verstynen, T. (2018). Population-averaged atlas of the macroscale human structural connectome and its network topology. *Neuroimage*, 178, 57-68. doi:10.1016/j.neuroimage.2018.05.027
- Yeh, F. C., & Tseng, W. Y. (2011). NTU-90: a high angular resolution brain atlas constructed by q-space diffeomorphic reconstruction. *Neuroimage*, 58(1), 91-99. doi:10.1016/j.neuroimage.2011.06.021
- Yeh, F. C., Verstynen, T. D., Wang, Y., Fernández-Miranda, J. C., & Tseng, W. Y. (2013). Deterministic diffusion fiber tracking improved by quantitative anisotropy. *PLoS One*, 8(11), e80713. doi:10.1371/journal.pone.0080713
- Yeh, F. C., Wedeen, V. J., & Tseng, W. Y. (2010). Generalized q-sampling imaging. *IEEE Trans Med Imaging*, 29(9), 1626-1635. doi:10.1109/tmi.2010.2045126
- Zhang, N., Gordon, M. L., & Goldberg, T. E. (2017). Cerebral blood flow measured by arterial spin labeling MRI at resting state in normal aging and Alzheimer's disease. *Neuroscience and biobehavioral reviews*, 72, 168-175. doi:10.1016/j.neubiorev.2016.11.023
- Zhang, W., Liu, J., Tian, L., Liu, Q., Fu, Y., & Garvey, W. T. (2013). TRIB3 mediates glucose-induced insulin resistance via a mechanism that requires the hexosamine biosynthetic pathway. *Diabetes*, 62(12), 4192-4200. doi:10.2337/db13-0312
- Zhang, W., Wu, M., Kim, T., Jariwala, R. H., Garvey, W. J., Luo, N., . . . Garvey, W. T. (2016). Skeletal Muscle TRIB3 Mediates Glucose Toxicity in Diabetes and High-Fat Diet-Induced Insulin Resistance. *Diabetes*, 65(8), 2380-2391. doi:10.2337/db16-0154
- Zhao, Z., Ho, L., Suh, J., Qin, W., Pyo, H., Pompl, P., . . . Pasinetti, G. M. (2003). A role of P301L tau mutant in anti-apoptotic gene expression, cell cycle and apoptosis. *Mol Cell Neurosci*, 24(2), 367-379. doi:10.1016/s1044-7431(03)00175-1
- Zheng, Y. Q., Zhang, Y., Yau, Y., Zeighami, Y., Larcher, K., Mistic, B., & Dagher, A. (2019). Local vulnerability and global connectivity jointly shape neurodegenerative disease propagation. *PLoS Biol*, 17(11), e3000495. doi:10.1371/journal.pbio.3000495
- Zlokovic, B. V. (2011). Neurovascular pathways to neurodegeneration in Alzheimer's disease and other disorders. *Nature reviews. Neuroscience*, 12(12), 723-738. doi:10.1038/nrn3114

# Chapter 3: Patient-centered transcriptomic and multimodal neuroimaging determinants of clinical progression, physical activity, and treatment needs in Parkinson's disease

## 3.1 Preamble

In the previous chapter, we developed a dynamical systems model and successfully applied it to study normal aging and AD. In this chapter, we examined the applicability of the model to PD. By incorporating the expression of a thousand genes with changes in dopamine, grey matter, and white matter, we identified the biological mechanisms underlying PD progression. We first confirmed if the model could successfully identify PD as the principal underlying neurodegenerative disease. Then using multiple clinical evaluations, we disentangled the distinct biological mechanisms underlying heterogeneity in clinical profiles. The model was also used to identify important biological pathways that mediate physical activity in PD. Finally, we demonstrated the translational utility of the model by using it to identify putative PD drugs through *in silico* gene perturbation.

This work is current under review in **npj Parkinson's Disease** journal as:

Quadri Adewale, Ahmed Faraz Khan, Sue-Jin Lin, Tobias R. Baumeister Yashar Zeighami Felix Carbonell, Daniel Ferreira, Yasser Iturria-Medina. "Patient-centered Transcriptomic and Multimodal Neuroimaging Determinants of Clinical Progression, Physical Activity and Treatment Needs in Parkinson's Disease".

## 3.2 Abstract

Parkinson's disease is a complex and multifactorial disorder, but how its biological and clinical complexity emerge from molecular to macroscopic brain interactions remains poorly understood. Here, we use a personalized multiscale generative brain model to characterize direct spatiotemporal links between genes and multimodal neuroimaging-derived biological factors in PD. We identified a set of genes modulating PD-associated longitudinal changes in dopamine transporter level, neuronal activity, dendrite density and tissue atrophy. Inter-individual heterogeneity in the gene-mediated biological mechanisms is associated with five distinct configurations of PD motor and non-motor symptoms. Although characterized by distinctive biological pathways, all the symptom configurations are associated with cell cycle processes. Notably, the protein-protein interaction networks underlying these configurations revealed distinct hub genes including *MYC*, *CCNA2*, *CCDK1*, *SRC*, *STAT3* and *PSMD4*. We also uncovered the biological mechanisms associated with physical activities performance in PD, and observed that leisure and work activities are principally related to neurotypical cholesterol homeostasis and inflammatory response processes, respectively. Finally, patient-tailored *in silico* gene perturbations revealed a set of putative disease-modifying drugs with potential to effectively treat PD, most of which are associated with dopamine reuptake and anti-inflammation. Our study constitutes the first self-contained multiscale approach providing comprehensive insights into the complex multifactorial pathogenesis of PD, unravelling key biological modulators of physical and clinical deterioration, and serving as a blueprint for optimum drug selection at personalized level.

## 3.3 Introduction

Parkinson's disease (PD) is a pervasive neurodegenerative disorder that presents with a variety of clinical manifestations such as motor (e.g., rigidity, resting tremor, bradykinesia),

psycho-cognitive (e.g., cognitive decline, depression, anxiety) and autonomic symptoms (e.g., constipation, hyposmia, sleep disorder). However, patients display heterogeneous combinations of symptoms, severity, and disease progression. Therefore, the etiopathogenesis of PD points to multiple probable causes including genetics, environment, and lifestyle (Simon et al., 2020). But the complex interplay between these biological factors is not clearly understood. Moreover, the pathological processes leading to the disease recruit many biological pathways at different cellular and molecular levels (Dong-Chen et al., 2023). Thus, a comprehensive framework incorporating several disease-associated variables is crucial for advancing the understanding of the disease. This is further supported by the recent efforts in transitioning towards a biological definition of PD (Höglinger et al., 2024; Simuni et al., 2024)

Current PD treatments are only symptomatic, and no single drug addresses the wide range of symptoms seen in patients. The mainstay of PD treatment, dopamine replacement therapy, relieves motor symptoms for a considerable number of patients, especially at the early stage of disease (Armstrong & Okun, 2020). However, 9-16% of patients do not respond to dopamine-based therapies, suggesting that patient heterogeneity plays a pivotal role in treatment response. Disappointingly, even hitherto responsive patients subsequently experience medication dose wear off and drug-associated worsening of symptoms, including drug-resistant tremor and medication-induced dyskinesias. As an adjuvant non-pharmacologic intervention, physical therapy helps improve a broad range of symptoms (Armstrong & Okun, 2020; Mak et al., 2017). However, the biological mechanisms underpinning the interaction of physical activity with PD are not fully understood. Furthermore, despite the perceived benefits of physical activity, sedentariness is still found among PD patients due to debilitating motor symptoms and other barriers such as perceived low expected benefit, lack of time, fear of falling, etc., that prevent patients from conducting exercise regimens (Ellis et al., 2013). Understanding the biological substrates of physical activity in PD can facilitate the discovery

of pharmacological alternatives, especially for those patients in advanced stages of the disease when physical activities are almost impractical.

Neuroimaging measures offer standard tools for routine clinical diagnosis of neurodegenerative diseases, further enabling the elucidation of disease pathogenesis and progression. Initial attempts integrating several neuroimaging modalities with either gene expression or receptor densities provided insights into the multiscale interactions in healthy aging and Alzheimer's disease (Adewale et al., 2021; Khan et al., 2021). This approach, called multifactorial causal modelling (MCM) (Iturria-Medina et al., 2017), affords a mechanistic way of understanding how the longitudinal changes in a biomarker emerge from the complex interplay between several biomarkers. Unified multimodal neuroimaging and expression of hundreds of genes revealed critical genetic determinants of healthy aging and Alzheimer's disease, as well as biological mechanisms separating the two processes (Adewale et al., 2021). The applicability of such unified multifactorial approach to subject-level modelling offers the unprecedented opportunity to harness inter-patient heterogeneity for better treatment plans and clinical trial design.

In this study, we extend the multiscale characterization of PD in four fundamental ways: (i) integrating whole-brain gene expression with longitudinal molecular, functional and (micro)structural neuroimaging-derived biological factors to infer gene-mediated brain reorganization in 89 PD patients from the Parkinson's Progression Markers Initiative (PPMI) cohort, (ii) linking different configurations of PD symptoms to distinct biological mechanisms and protein-protein interaction networks, (iii) identifying molecular mediators of the interplay between PD progression and physical activity, and (iv) using patient-level *in silico* gene perturbations to identify putative disease-modifying drugs for PD. This work represents a pioneering attempt to unify multiple aspects of PD-associated biomarkers and physical

activities at different resolutions, paving the way for a deeper understanding of PD biological mechanisms and identifying effective personalized treatments.

## 3.4 Results

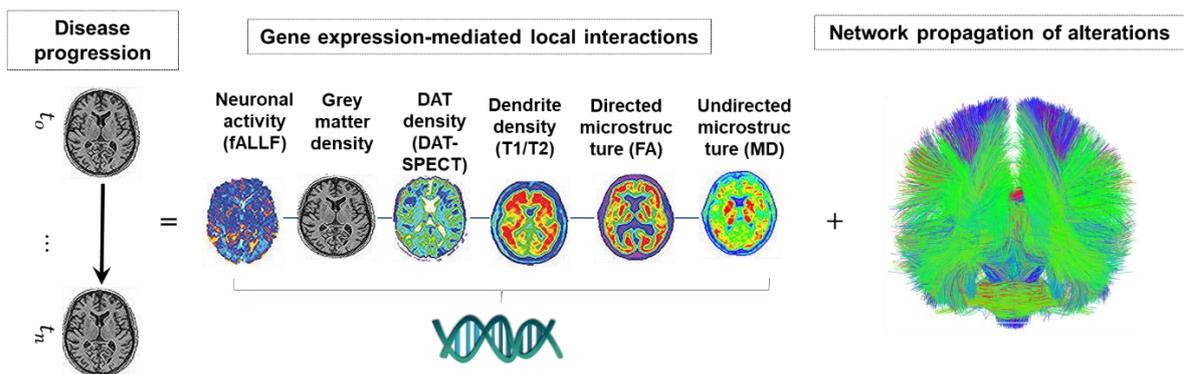
### 3.4.1 Whole-brain multiscale transcriptomic-neuroimaging model of Parkinson's disease

To characterize widespread molecular, functional and structural brain changes in PD patients at the individual level, we fit a whole brain model with gene expression and six longitudinal neuroimaging-derived biological factors. These imaging modalities macroscopically capture typical neurodegenerative changes, namely, dopaminergic loss (DAT-SPECT), neuronal activity (fALFF), directed microstructural changes (fractional anisotropy), undirected microstructural damage (mean diffusivity), dendrite density (t1/t2 ratio (Righart et al., 2017)), and atrophy (gray matter density). They are acquired over multiple scans in 89 PD patients from the PPMI cohort. The transcriptomic data was derived from 6 neurotypical brains from the Allen Human Brain Atlas (AHBA) (Hawrylycz et al., 2012) across 976 landmark genes, which have been shown to be central to biological functions and recapitulate about 89% of the whole human transcriptome (Subramanian et al., 2017). Anatomical connectivity was estimated from the high-resolution Human Connectome Project template (HCP-1065; *Methods: Anatomical connectivity estimation*).

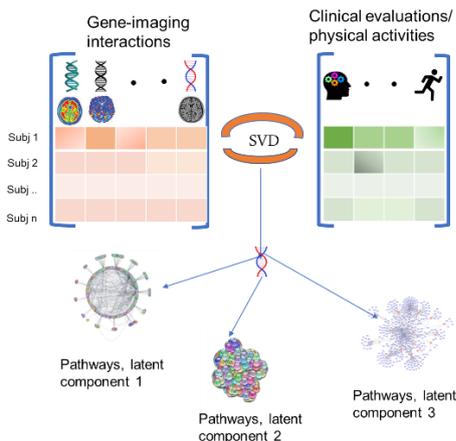
Our mathematical framework, named gene expression multifactorial causal model (GE-MCM; Figure 3.1A), is formulated to capture the influence of gene expression on a particular biological factor and accounts for the network-mediated spreading of the subsequent aberrant changes across the brain (see *Methods: Gene Expression Multifactorial Causal Model*). Using a robust Bayesian optimization technique, we estimated regression coefficients (gene-imaging parameters) that capture the modulation effect of each gene on the dynamic changes and interactions of the individual imaging derived biological factor. Even though we used a single

fixed neurotypical gene expression template across the population (AHBA), the personalized gene-imaging parameters quantify individual gene dysregulation patterns and serve as proxies for gene-specific deviations needed for individual model fitting. Indeed, when applied to the studied PD population, the model showed a good predictive ability to reproduce the six disease-affected longitudinal imaging-derived biological factors ( $R^2 = 0.71 \pm 0.2$ ). Notably, the model parameters demonstrated the capacity to (i) correctly unravel the biological mechanisms underlying inter-patient variability in clinical manifestations or physical activity (Figure 3.1B) and (ii) infer patient-specific complete model for *in silico* drug discovery via gene perturbation (Figure 3.1C).

### A Transcriptome-mediated multifactorial causal model of interacting biological factors



### B Association of gene-imaging interactions with clinical symptoms and physical activities



### C Perturbation of gene expression for drug discovery

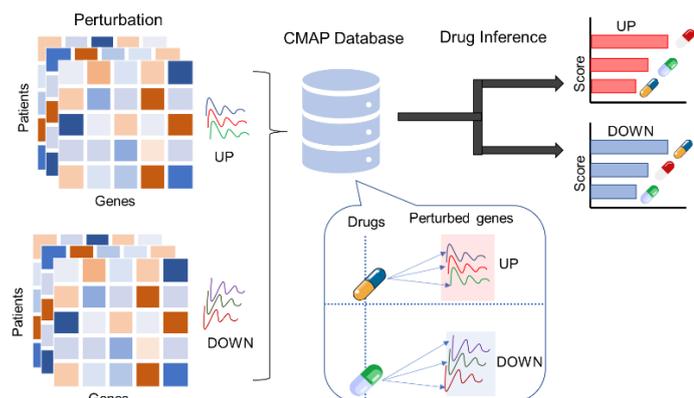


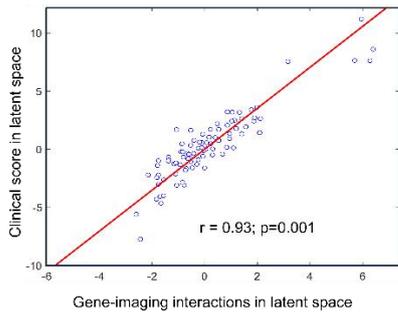
Figure 3.1: Transcriptomic-neuroimaging multifactorial causal modeling of PD. A) Patient's temporal disease evolution captured by multimodal neuroimaging is decomposed into i) local transcriptome-mediated interactions between neuroimaging measures disease-related biological factors, namely dopaminergic loss, neuronal activity, directed microstructural changes, undirected microstructural damage, dendrite density, and neuronal atrophy ii)

network-mediate propagation of pathological effects between brain regions. The patient-specific gene-imaging parameters  $\{\alpha\}$  are obtained by robust Bayesian regressions optimizing the differential equations (Methods). B) Covariance between the gene-imaging parameters and slopes of clinical evaluations or physical activities are resolved along multiple principal axes to unveil the underlying biological pathways. C) In silico bidirectional perturbation of genes identifies putative PD drugs. The perturbation of a therapeutic gene is expected to cause a slower disease progression when compared to progression without perturbation.

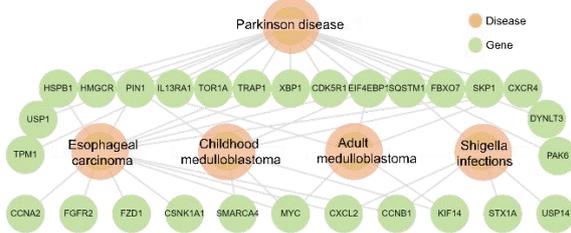
### 3.4.2 Identifying transcriptomic mechanisms mediating behavioural and cognitive deterioration in PD

We sought to identify genetic drivers of multifactorial brain reorganization due to PD progression, particularly those genes controlling direct spatiotemporal interactions among dopaminergic loss, neuronal activity, directed microstructural changes, undirected microstructural damage, dendrite density, and atrophy. First, out of a total of 35,136 gene-imaging parameters, we identified 953 stable parameters whose 95% confidence intervals (CI) exclude zero. Singular value decomposition (SVD) was then used to find the shared latent space between these stable parameters and 11 different clinical evaluations (*Methods: Clinical and Physical Activity Measures*). Five of the eleven principal components are significant following permutation tests ( $p < 0.05$ ). However, the first principal component accounts for a notable proportion (43.7%) of the explained covariance. Projection of the gene-imaging interactions and clinical evaluations on this first latent component showed a very high correlation of  $r = 0.93$  ( $p = 0.001$ ; Figure 3.2A). Furthermore, we discovered 85 genes with significant contributions to the axis (bootstrap ratio  $> 1.96$ ). Interestingly, querying the diseases associated with the genes in DisGeNET database revealed PD as the leading disorder ( $q < 0.05$ ; Figure 3.2B). The identification of esophageal carcinoma, medulloblastoma and shigella disease corroborates bodies of evidence associating cancers and gut disorders with PD. Shigella and Escherichia coli are major causes of diarrhea, and Shiga toxins is linked to damage in blood-brain barrier, microvasculature, astrocytes and neuron with characteristic motor symptoms (Pinto et al., 2017). Similarly,  $\alpha$ -synuclein (*SNCA*) has been suggested as a biomarker for medulloblastoma (Y.-X. Li et al. 2018). The results support the relevance of the identified genes to PD pathogenesis and its systemic interaction with other disorders.

**A Correlation between gene-imaging interactions and clinical symptoms in Parkinson's disease**



**B Diseases associated with the genes driving clinical symptoms**



**C Multiple genes interact with several biological factors to drive Parkinson pathology**

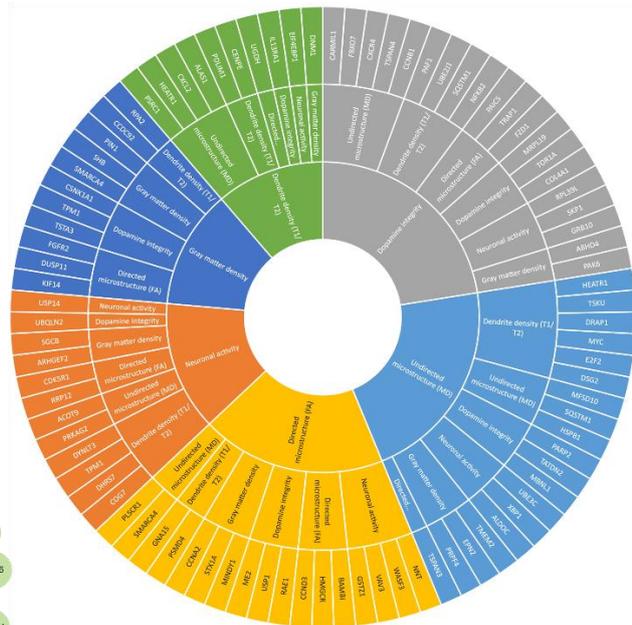


Figure 3.2: Genes underlying clinically-relevant spatiotemporal vulnerability in PD overlap with cancer and infection pathways. **A**) High correlation ( $r=0.93$ ;  $p=0.001$ , FWE-corrected) between projections of gene-imaging pathological interactions and rates of clinical deterioration on the first latent component of SVD. The first latent component accounts for 43.7% ( $p=9.99 \times 10^{-4}$ , FWE-corrected) of the explained covariance between clinical evaluations and the pathological interactions of gene expression and neuroimaging measures. **B**) DisGeNET disease-gene associations of the genes contributing to the pathological interactions on the first latent component. The hierarchical relationship shows the significant cognate diseases and their shared genes ( $q < 0.05$ ). **C**) Multifactorial interactions between identified genes and neuroimaging-derived biological factors. A gene directly influences how a neuroimaging-quantified biological factor interacts with other factors to cause a factorial alteration along the disease's course. Notably, the outermost ring represents the genes modulating the interactions among biological factors, the middle ring displays the biological factor directly influenced by a gene, and the innermost ring shows the biological factor undergoing longitudinal changes because of the interactions.

We further investigated the specific structural, functional, microstructural, or dopaminergic changes that are modulated by the 85 genes. Since each optimized gene-imaging parameter associates a gene with a biological factor, we retrieved the biological factors of the significant parameters associated with the 85 genes. We observed a broad range of interactions between the genes and the six disease-related biological factors (Figure 3.2A). Among the PD-related genes, we observed that *TPM1* modulates dopamine level in driving longitudinal changes in atrophy, which is consistent with the gene's role in controlling striatal dopamine release (Downs et al., 2021; Wakabayashi-Ito et al., 2011). Our results further suggested that *TPM1* also modulates dendrite density to drive longitudinal alterations in neuronal activity, in

agreement with the gene's activity of regulating actin filament and neurite growth (Brettle et al., 2016). Similarly, we found that *CXCR4* modulates mean diffusivity (a measure of myelin or axon integrity) to drive dopamine change longitudinally. Activation of *CXCR4* promotes the development of oligodendrocytes for remyelination of injured adult central nervous system (Patel et al., 2010). Overall, our findings transcend traditional single-scale transcriptomic or neuroimaging analysis by considering biologically plausible complex interactions underlying PD progression.

### **3.4.3 Uncovering the protein-protein interaction networks underlying PD phenotypic landscapes**

To understand how the model-derived pathological interactions might be related to the different clinical manifestations of PD, we analysed all the five significant latent components of the SVD. The explained co-variance of these components are 43.7%, 14.5%, 10.2%, 7.1% and 6.3%, respectively. Projecting the 11 clinical scores onto these components allowed us to disentangle the contributions of psychiatric, motor, cognitive and other PD symptoms to each latent component. Using a high confidence score (cut-off=0.7), we then retrieved the protein-protein interaction (PPI) networks of the genes associated with each component from STRING database (Szklarczyk et al., 2021). The biological pathways ( $q < 0.05$ ) relevant to PPI networks were also obtained from Wikipathways.

In contrast to other components, the first latent component, whose genes were earlier associated with PD in Figure 3.2, shows a balanced contribution from the four groups of symptoms (Figure 3.3A). This observation indicates that the leading biological mechanisms underlying PD engender a wide range of clinical symptoms. Nevertheless, the largest individual symptomatic contribution comes from motor signs (UPDRS-III), the principal hallmark of PD. The associated PPI network points to the active roles of cell cycle, DNA damage, and rapamycin signaling. The second component is dominated by cognitive

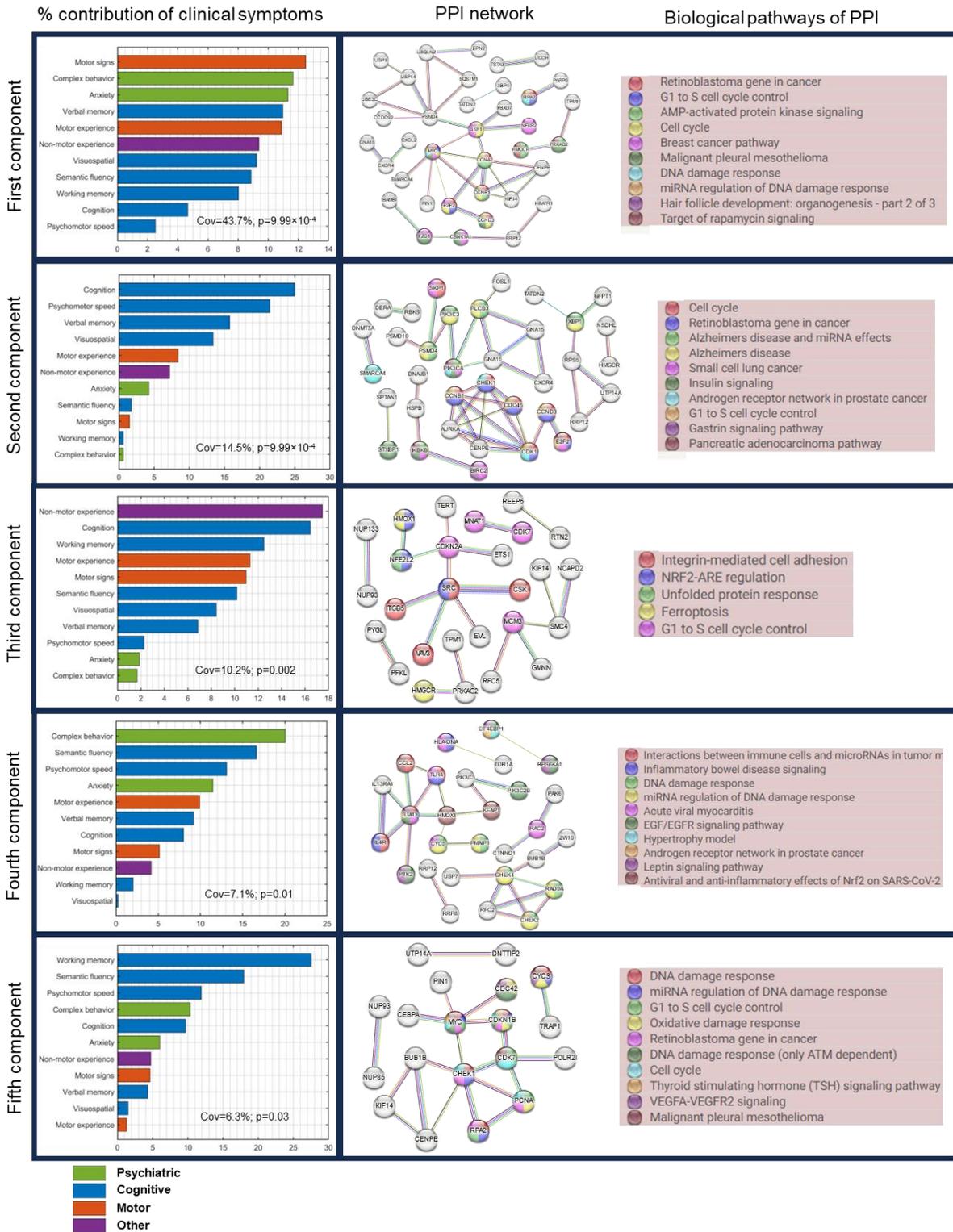


Figure 3.3: Associations of patterns of PD clinical symptoms with biological mechanisms. *Left:* Five significant latent components were identified based on permutation analysis ( $FWER < 0.05$ ) of shared covariance between gene-imaging pathological interactions and clinical evaluations. Bar graph shows the relative contributions of each clinical evaluations, grouped by symptom types, on each latent component. *Middle:* Protein-protein interaction networks (PPI) of significant genes associated (bootstrap ratio  $> 1.96$ ) with each latent component. PPPI networks were retrieved from STRING with a high confidence score  $\geq 0.7$ . *Right:* Top biological pathways ( $q < 0.05$ ) associated with the genes in the PPI networks. Colored nodes in the PPI networks correspond to genes implicated in the top biological pathways.

symptoms, which supports why the biological pathways include Alzheimer's disease, in addition to insulin signalling, cell cycle, and gastrin signalling. The largest contributions to the third component come from motor, cognitive and other non-motor symptoms of daily living (e.g., pain, fatigue, and autonomic dysfunctions.). Notably, the implicated pathways include ferroptosis, unfolded protein response, cell cycle and oxidative stress. The fourth and fifth components are predominantly psychiatric and motor symptoms, with suggested roles of inflammation, leptin signalling, cell cycle, DNA damage response, and oxidative stress. Despite the varied symptom profiles and underlying PPI networks, we observe a common association of cell cycle processes with all the latent components.

As hub genes are believed to play central roles in biological processes and gene regulatory networks (Yu et al., 2017), we sought to identify the leading hub genes in the PPI networks. Interestingly, each latent component has at least one dense PPI sub-network which could be prioritized for biomarker or drug discovery. We therefore selected the hub genes as those with the highest node degrees. We identified 3 hubs genes, namely, *MYC*, *CCNA2* and *PSMD4* in the first component due to a tie in their rankings. *CDK1*, *SRC*, and *STAT3* were ranked highest for the second, third and fourth components, respectively. The PPI network of the of fifth component was not queried because its enrichment PPI value was not significant ( $p = 0.194$ ; Figure 3.3). Apart from *PSMD4*, other genes have been previously identified as hub genes in PD. Our results however suggests that different hub genes might be associated with different patterns of clinical symptoms in PD.

#### **3.4.4 Molecular pathways associated with physical activity in PD**

Physical activity reduces the risk of developing PD and ameliorates s both motor and non-motor PD symptoms (Amara et al., 2019; Langeskov-Christensen et al., 2024; Paul et al., 2019). Conversely, the symptomatology of PD presents many barriers (such as motor dysfunction, cognitive impairment, depression and apathy) to engaging in physical activities,

(Amara et al. 2019). Molecular pathways modulating the relationship between PD and physical activity may therefore shed light onto key neuroprotective mechanisms. We therefore investigated possible biological mechanisms associated with physical activity in PD by applying SVD to identify axes of covariance between the stable gene-neuroimaging parameters and three different domains of physical activity, namely household, work, and leisure activities. The individual scores for the different domains were derived from PASE, a self-reported questionnaire commonly used to quantify physical activity levels in older adults (Washburn et al., 1999). Two SVD principal components were relevant based on permutation tests ( $p < 0.05$ ), and they separately explained 47% and 37% of the data covariance. Leisure activities (e.g., resistance training, jogging, swimming) account for about half (49%) of the first axis (Figure 4). Conversely, work-related activities (e.g., walking and lifting) contribute (54%) principally to the second axis (Figures 3.4C). Nevertheless, household activity account for 27% and 40% of first and second axes, respectively.

Next, using the genes with significant contributions to each axis (bootstrap ratio  $> 1.96$ ), we queried the associated biological pathways from Reactome and WikiPathways. The two pathway databases were combined to ensure the robustness of our findings and avoid database bias. The first component is principally associated with cholesterol biosynthesis ( $q < 0.05$ ) while the second component is largely implicated in immune-related processes such as toll-like receptor and B cell signalling. Even though reverse causation cannot be disregarded (as reduced activity levels may accelerate PD progression, and vice versa), the identified biological pathways may be partly explaining individual predisposition/variation to physical activity under PD effects.

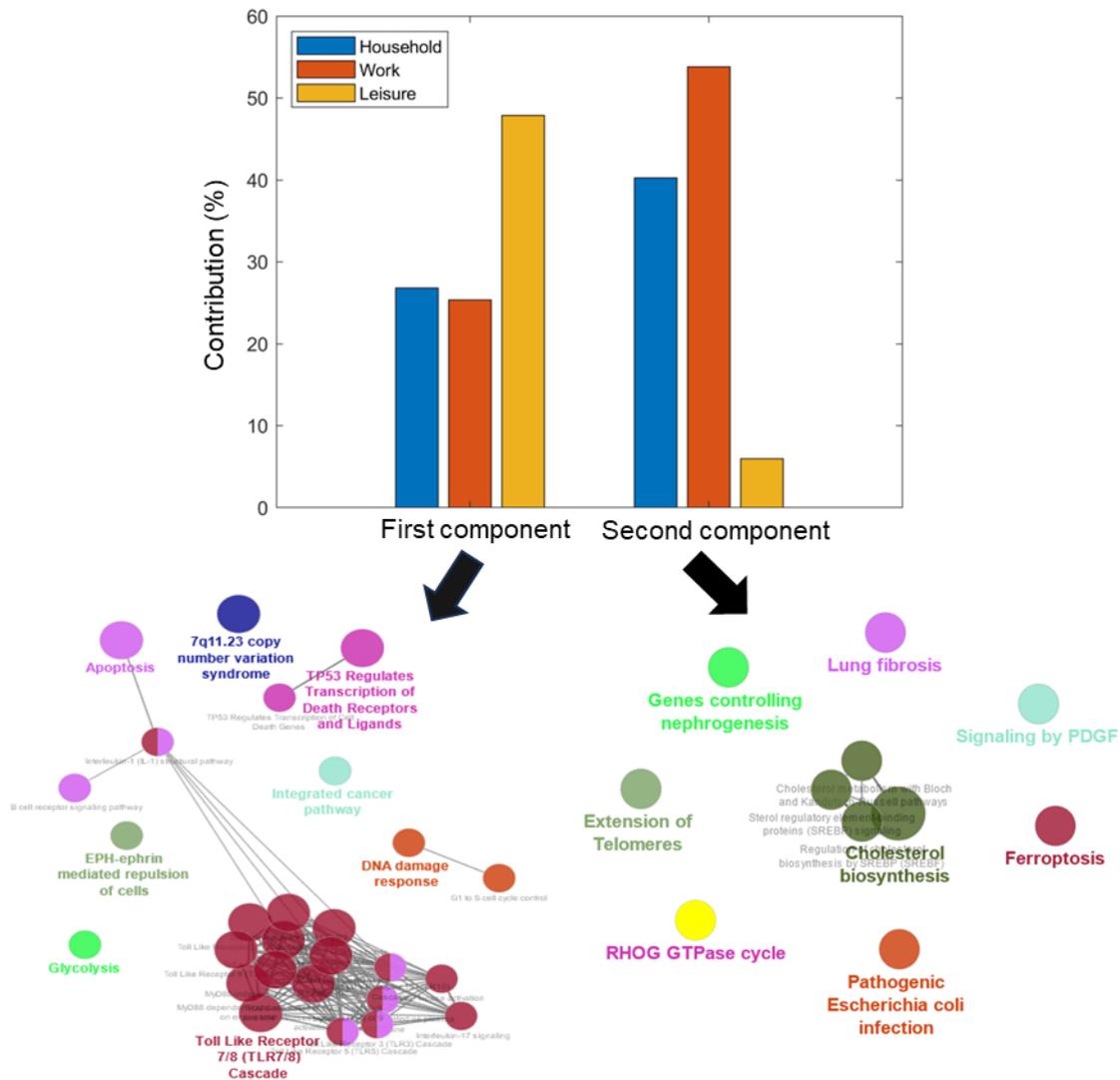


Figure 3.4 : Distinct molecular pathways are associated with leisure and work activities. Contributions of the three domains of the Physical Activity Scale for the Elderly (PASE) towards the first and second principal axes. The corresponding biological pathways mediating the interactions between physical activity and PD in each axis are pointed by the arrows. The primary and secondary axes were obtained from the singular value decomposition (SVD) of the covariance matrix of the gene-neuroimaging parameters and slopes of physical activity scores. The relevant biological pathways were queried from Reactome and Wikipathways ( $q < 0.05$ ).

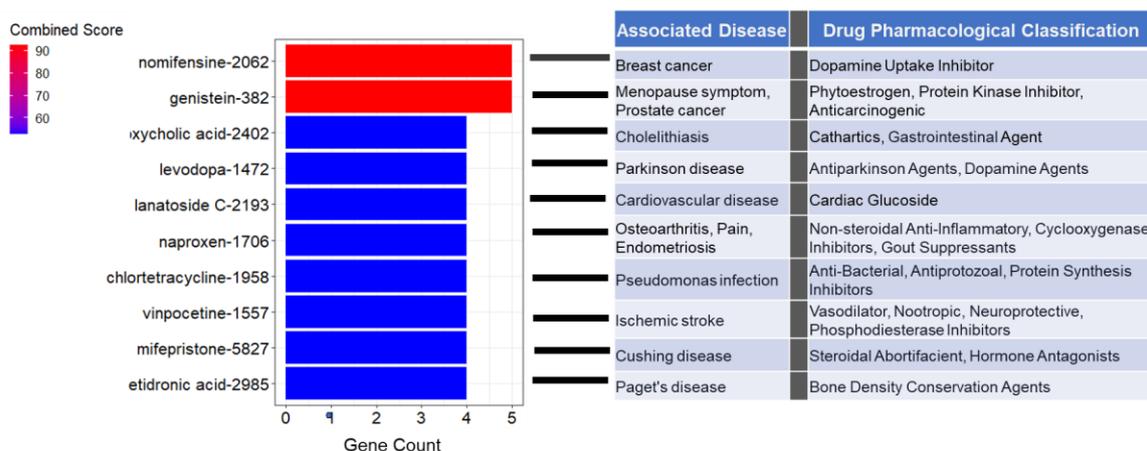
### 3.4.5 Virtual Gene Perturbations Reveal Potentially Effective Drug Candidates

Finally, we used a multifactorial-causal perspective to predict potential therapeutic drugs candidates for PD. For this, we utilized the individually fitted GE-MCM to simulate the disease's subsequently progression for 2 years after the last evaluated time point. We then up- and down-perturbed each gene and quantified the influence of each perturbation on the brain's multiregional and multifactorial imaging descriptors associated with disease progression

(*Methods: Gene Perturbation for Drug Discovery*). A gene was considered therapeutic if the perturbation-induced brain changes implied a slower disease progression than the actually observed within the two years under consideration. We then ranked the genes based on the number of subjects for which they have a therapeutic effect and selected the top genes in the 90<sup>th</sup> percentile. Next, using the CMap database in EnrichR, we queried the inverse-drug relationship between selected genes and several drugs. CMap allowed us to map previous drug-induced transcriptomic perturbations to our *in silico* perturbation profiles. We checked for the alignment between the genes up- and down-regulated by drugs in CMap and our up- and down-perturbed therapeutic genes, respectively. We then retrieved the associated disease and pharmacological classification of these drugs from PubChem.

Figures 3.5A and 3.5B show the list of the top respective drug candidates ranked by combined score (product of odds ratio and negative natural log of the p-value). Three of the drugs are associated with dopamine, the principal neurotransmitter implicated in PD. The first among the list of drugs associated with the upwardly perturbed genes is nomifensine, a drug that increases synaptic dopamine availability by inhibiting dopamine reuptake (Figure 3.5A). Similarly, the fourth drug is Levodopa, the most commonly used drug for treating PD symptoms. Among the top drugs identified through the *in silico* down-perturbation is pergolide, an ergoline-based dopamine receptor agonist still being used to treat PD in some countries. Furthermore, we found a notable number of drugs currently used to treat infections, hinting at the potential of repurposing anti-infectives for PD treatment. Other drugs are implicated in cardiovascular disease, insomnia and inflammation. In sum, the identification of some of the current dopamine-base PD drugs demonstrates the potential of our *in silico* perturbation method to discover prospective drugs for PD treatment.

### A Top drugs associated with up-perturbed genes



### B Top drugs associated with down-perturbed genes

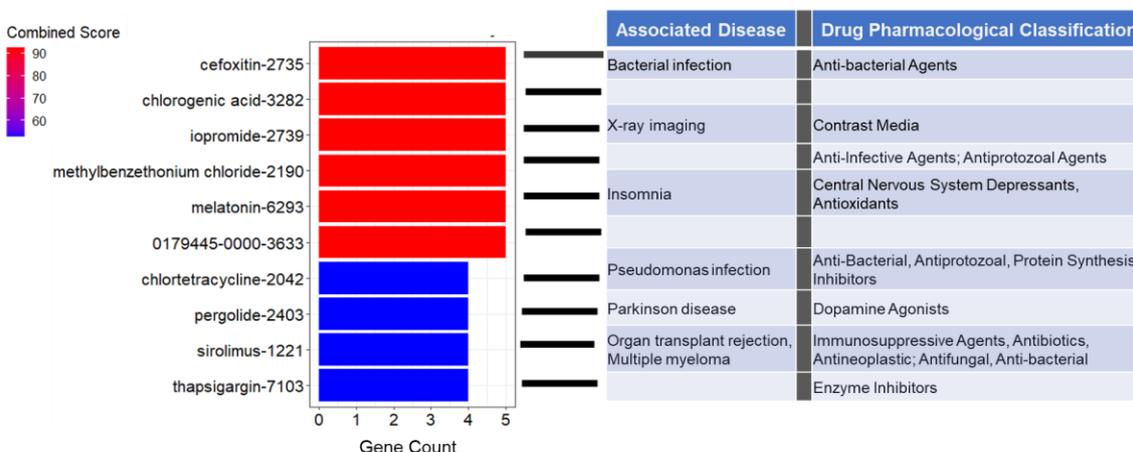


Figure 3.5 : Top drugs for PD identified by our virtual perturbation framework. A) The top 10 drugs identified by upward perturbation of genes. The therapeutic genes were discovered by increasing their expressions by 10% and observing the effect of disease progression within 2 years. The drugs were obtained by comparing the identified therapeutic genes with the transcriptomic effect of drugs from CMAP database. The drugs are ranked by the combined score (odds ratio  $\times$   $-\log(p\text{-value})$ ). B) The top 10 drugs identified by downward perturbation of genes.

## 3.5 Discussion

Parkinson's disease (PD) is a highly complex and heterogenous disease involving various biological mechanisms. We developed a novel computational approach that incorporates multimodal neuroimaging data, averaged template of bulk gene expression, clinical evaluations, and physical activities to unravel the multifactorial changes accompanying the disease process. We validated the relevance and informativeness of the personalized models by identifying PD as the primary neurodegenerative disease associated with the molecular

pathways and clinical symptoms. We further unravelled the biological substrates underpinning the relevance of physical activity to the disease course. Finally, we demonstrated the usefulness of our approach for drug discovery and repurposing via *in silico* transcriptomic perturbations. This first of its kind study presents a self-contained bottom-up causal approach for advancing the understanding of complex multilevel disease processes and identifying potential disease-modifying therapeutic targets.

The roles of genes in maintaining healthy aging and contributing to neurodegenerative diseases are not completely known. In this work, we modelled mechanistic interactions between imaging-derived longitudinal biological factors and spatial variability in gene expression. This modelling approach allowed us to uncover PD-relevant genes and the biological processes they interact with. For instance, among the genes identified are *PINI*, *SKP1*, *TRAP1*, *TORIA* (Figure 3.2A). *PINI* is expressed in neurons and found play active roles in neuronal cell death and apoptosis (Ghosh et al., 2013; Zhang et al., 2022). Mice lacking *PINI* displayed neuronal degeneration including motor and behavioral dysfunctions (Liou et al., 2003). Concordantly, our result suggests that *PINI* directly modulates gray matter density in PD (Figure 3.2A). Similarly, decreased expression of *SKP1* homolog (*SKPA*) and *TRAP1* have been shown to cause loss of dopamine in flies, accompanied by motor symptoms (Butler et al., 2012; Dabool et al., 2020). We found that the two genes directly interact with dopamine to drive longitudinal change in neuronal activity and directed microstructure, respectively. *TORIA* is highly expressed in the substantia nigra, a key region in the pathogenesis of PD, and is responsible for primary hereditary form of dystonia, partly due to its effect on striatal dopamine (Wakabayashi-Ito et al., 2011). Concordantly, our findings suggest that *TORIA* modulates dopamine to drive the longitudinal alterations in dopamine integrity. Overall, many of the gene-imaging relationships identified in our study of PD patients have also been reported *in vivo* in animal models. Hence, the novel gene-neuroimaging associations can be further

validated through experimental models. The insights afforded by these relationships can advance our mechanistic understanding of the disease and help to streamline the identification of possible off-targets when targeting genes for drug development.

Although the hallmark signs of PD are motor complications, dysregulation of multiple clinical domains including cognition, memory, mood, behavior, and autonomic functions supports the complex and multisystem view of the disease. Moreover, heterogeneity in patients' symptoms and response to treatment has led to the definition of various PD subtypes (Mestre et al., 2021). Even though our study did not subtype patients because of the small sample size, we uncovered five distinct axes of association between biological mechanisms of PD and clinical symptoms. Interestingly, we observed qualitative differences in the relative involvement of symptom types to these axes. Network biology approach also revealed different PPI topology and biological pathways underlying these symptom distribution profiles. Despite the association of numerous pathological processes such as protein aggregation, oxidative stress, ferroptosis, and neuroinflammation with PD, the link between these processes and heterogenous symptom manifestations are lacking. Our study aligns symptom profiles with biological pathways. We found that inflammation, leptin signalling DNA damage response and oxidative stress may be associated with predominant motor and psychiatric symptoms while insulin and gastrin signalling could be implicated in pronounced cognitive symptoms. Nevertheless, we observed a general involvement of G1/S cell cycle control (or its associated processes) with all the symptom distributions. Cell cycle re-entry in post-mitotic neurons might cause neurodegeneration by triggering response to oxidative stress, DNA damage, and other pathological processes (R. Sharma et al., 2017). Concordantly, rotenone-based model of PD showed that lowering the amount of rotenone reduces endoreplication-induced neurodegeneration by blocking cell cycle progression at G1/S phase (Frade & López-Sánchez, 2010; H. Wang et al., 2014). We further prioritized 6 hub genes related to these symptom

profiles. Interestingly, five of these genes (*MYC*, *CCNA2*, *CCDK1*, *SRC* and *STAT3*) have been previously identified as PD hub genes from different cohort studies of gene expression in substantia nigra and peripheral blood (Banerjee et al., 2021; Elango et al., 2023; George et al., 2019; Liu et al., 2019; M. Wang et al., 2023). The novel hub gene, *PSMD4*, is a receptor of the 26S proteasome which is responsible for protein degradation (Collins & Goldberg, 2017). Given the relevance of proteasome homeostasis to intracellular accumulation of  $\alpha$ -synuclein (Bi et al., 2021), this novel hub gene may play a key role in PD pathogenesis. The hub genes in this study can guide the identification of druggable targets and biomarkers for heterogeneous PD symptom profiles.

The benefits of physical activity to PD symptoms and progression are widely acknowledged. Even though the biological mechanisms mediating these benefits are fully understood, physical activity may promote neuronal plasticity and survival of dopaminergic neurons by simulating the expression of neural growth factors (Da Silva et al., 2016). Here, we found that physical activity is associated with PD through two principal pathways, namely, cholesterol biosynthesis and inflammation via toll-like receptors. A previous study of animal model of PD showed that MPTP-bearing mouse had reduced  $\alpha$ -synuclein and downregulation of toll-like receptors after eight weeks of treadmill exercise (Koo et al., 2017). Although the results on the association of cholesterol with PD are mixed, several PD-related genes are involved in cholesterol homeostasis (García-Sanz et al., 2021; Jin et al., 2019). Moreover, cholesterol biosynthesis has been shown to decrease in the fibroblasts of PD patients (Musanti et al., 1993). The most compelling insights into the tripartite association between PD, cholesterol and physical activity was demonstrated recently (Dutta et al., 2022). The authors found that physical activity activates PPAR $\alpha$  in the dopaminergic neurons of PD mouse model. Activation of PPAR $\alpha$  alone suppressed the aggregation and spreading of  $\alpha$ -synuclein in the mouse. As PPAR $\alpha$  is a transcription factor that regulates the expression of genes involved in

fatty acid oxidation, the mouse was treated with fenofibrate, a PPAR $\alpha$  medication for abnormal cholesterol level. The authors observed that one month of daily treatment with fenofibrate conferred similar benefits as two months of regular exercise. Despite that our analysis does not rule out the bidirectional relationship between PD and physical activity, our results are consistent with the foregoing studies. However, the mode and intensity of exercise remains an open question. A meta-analysis of 19 randomized human clinical trials showed that different modes and regimens of exercise provide different forms of benefits to PD symptoms (Tang et al., 2019). Indeed, our findings could guide a more personalized prescription of physical activity in PD. Perhaps, leisure-related activities (likely shorter duration, higher intensity) would be more beneficial to patients having abnormal cholesterol levels while home- or work-related activities (likely repetitive and lower intensity) could help with neuroinflammation-induced PD pathogenesis. Furthermore, personalized physical activity regimen can be prescribed by comparing the gene-neuroimaging parameters of a patient with the parameters of other patients who have benefited from a particular exercise regimen.

Current treatments for PD are symptomatic, hence the search for disease-modifying treatments addressing the underlying pathology is a priority. While the mainstay of PD treatment are dopamine-based drugs, their effectiveness largely varies with disease subtype and stage (Armstrong & Okun, 2020). Interestingly, among the top 20 putative PD drugs identified in our study, there are three dopamine-based drugs, including levodopa, the current first line treatment for PD (Figures 3.4A and B). However, we also identified multiple immune-related and anti-inflammatory drugs, including naproxen (a non-steroidal anti-inflammatory drug) and tetracycline. Other drugs such as vinpocetine, chlorogenic acid and melatonin have also been reported to modulate inflammation. Although vinpocetine is typically prescribed for treating memory loss in aging and dementias (including PD patients with dementia), it has been demonstrated to regulate the circulation of inflammatory molecules in PD patients (Ping et al.,

2019). Chlorogenic acid, found in coffee, is suggested to offer neuroprotective roles in animal models of PD (He et al., 2021; N. Sharma et al., 2022; Singh et al., 2018). This latter finding may partially explain why coffee confers reduced risk on PD development. Similarly, melatonin, which may improve sleep disturbance in PD (Srinivasan et al., 2011), has also been shown to reduce neuroinflammation (Li et al., 2022). The convergence of these medications on immune system/inflammation highlights the need to consider this pathway for drug discovery and repurposing. Even though we performed our drug query using therapeutic genes across the patient population, personalised treatment can be designed by querying the drug database with patient-level therapeutic genes.

The lack of patient-specific gene expression data constrained us to use a single neurotypical gene expression template. Nevertheless, we previously demonstrated that the interaction of the static transcriptomic information with patient-specific longitudinal neuroimaging measures provides a proxy for patient-specific genetic deformation in healthy aging and Alzheimer's disease (Adewale et al., 2021). Furthermore, the static gene expression data was obtained by combining the mRNA values of six different subjects and inferring the gene expression for the brain regions with missing values (Adewale et al., 2021). Despite the inherent variability and bias that could arise from inter-subject variability and mRNA interpolation, the identification of PD as the underlying neurodegenerative disease demonstrates the validity of our approach (Figure 3.2B). Subject-specific gene expression may help refine the derived gene-imaging parameters and better facilitate personalized treatments.

Overall, our universal mathematical formulation can be used to study other multifactorial and progressive disorders such as frontotemporal dementias and amyotrophic lateral sclerosis. As subtyping often requires a large number of subjects and raises a question of within-subtype homogeneity, the gene-imaging parameters provide a way to mechanistically capture biological and clinical variability for better treatment plans in heterogenous diseases.

Future work will also consider how these parameters can predict patient response to treatment in clinical trials.

## **3.6 Methods**

### **3.6.1 Ethics statement**

This article does not contain any studies with human participants performed by any of the authors. The neuroimaging and clinical data were acquired from the multicenter Parkinson's Progression Markers Initiative (PPMI; [ppmi-info.org](http://ppmi-info.org)). As per PPMI protocols, study participants and/or authorized representatives gave written informed consent at the time of enrollment for sample collection and completed questionnaires approved by each participating site Institutional Review Board (IRB). The authors obtained approval from the PPMI for data use and publication, see documents <https://www.ppmi-info.org/documents/ppmi-data-use-agreement.pdf> and <https://www.ppmi-info.org/documents/ppmi-publication-policy.pdf>, respectively.

### **3.6.2 Data description and processing**

#### ***Study participants***

This study involved 89 individuals from PPMI (RRID:SCR\_006431) (<http://ppmi-info.org/>). The subjects have at least three imaging modalities out of the following: structural MRI, resting functional MRI, diffusion MRI, dopamine SPECT; for at least three visits. The PPMI was launched in 2010 as an observational study of longitudinal changes in volunteer subjects with and without PD. PPMI is led by Principal Investigator Kenneth Marek, MD and sponsored by the Michael J. Fox Foundation, with the goal of understanding the onset and progression of PD.

## ***Structural MRI***

Structural T1- and T2-weighted 3D brain images were acquired as described in PPMI manuals (<http://www.ppmi-info.org/>). The images were corrected for intensity nonuniformity using the N3 algorithm (Sled et al., 1998). They were segmented into grey matter (GM), white matter (WM), and cerebrospinal fluid (CSF) probabilistic maps, using SPM12 (<http://www.fil.ion.ucl.ac.uk/spm>). The gray matter segmentations were standardized to MNI space (Evans et al., 1994) using DARTEL (Ashburner, 2007). Each map was corrected for the effects of spatial registration to preserve the initial amount of tissue volume. Mean gray matter density values of the T1- and T2-weighted images were calculated for a total of 163 grey matter regions described in *Methods: Gene Expression and Brain Parcellation*.

## ***Resting-state fMRI***

Resting-state functional images were acquired using an echo-planar pulse sequence on a 3.0T Philips MRI scanner with the following parameters: 140 time points, repetition time (TR) = 2400 ms, echo time (TE) = 25 ms, flip angle = 80°, number of slices = 40, slice thickness = 3.3 mm, in-plane resolution = 3.3 mm, and in-plane matrix size = 68 × 66. The fMRI images were preprocessed using FSL (v5.0) toolbox (<https://fsl.fmrib.ox.ac.uk/fsl/fslwiki>) (S. M. Smith et al., 2004). The preprocessing steps are: 1) Motion and slice timing correction 2) Alignment to the structural T1 image 3) Spatial normalization to the MNI space using the registration parameters obtained for the structural T1 image with the nearest acquisition date, and 4) Signal filtering to retain only low-frequency fluctuations (0.01–0.08 Hz) (Chao-Gan and Yu-Feng, 2010). Due to its high sensitivity to disease progression (Iturria-Medina et al., 2016), we used fractional amplitude of low-frequency fluctuation (fALFF) as a regional quantitative indicator of the brain's functional integrity. fALFF quantifies resting-state regional brain activity as the ratio of the power spectrum of the low frequency band (0.01 – 0.08 Hz) to the power spectrum of the whole frequency range (0 - 0.25Hz) (Zou et al., 2008).

## ***Diffusion MRI***

Diffusion MRI (dMRI) was obtained using standardized protocols on Siemens Verio and Siemens Tim Trio 3T MRI scanners. A single-shot echo-planar imaging scheme was used with 64 sampling directions, a b-value of 1000 s/mm<sup>2</sup> and a single b = 0 image. Other parameters include 116 × 116 matrix, 2 mm isotropic resolution, TR/TE 900/88 ms, and two-fold acceleration. More information on the dMRI acquisition and processing can be found online at <http://www.ppmi-info.org/>. Further preprocessing was done in FSL (v5.0). First, the DTI scans were corrected for motion, eddy current and EPI distortion. Then, the b0 images were aligned to the corresponding subject's T1-weighted images based on mutual information. The deformation field between the diffusion and T1-weighted image was calculated. The deformation field and eddy current transformations were applied to the dMRI images. Diffusion tensor models were then fitted independently for each voxel. Next, the scans were normalized to MNI space (Evans et al., 1994) using the registration parameters obtained for the structural T1 image with the nearest acquisition date. The mean values of the fractional anisotropy and mean diffusivity were estimated for each of the 163 brain regions of interest.

## ***Dopamine SPECT***

A 111-185 MBq (3-5 mCi) bolus injection of I-123 FB-CIT was administered to each participant and the SPECT scans were obtained 4 hours post-injection. Raw projection data was acquired as a 128x128 matrix, after which the SPECT image was reconstructed. The images were preprocessed using SPM12. The scans underwent for attenuation correction and noise reduction using Gaussian blurring with a 3D 6mm filter were applied. The reconstructed and corrected SPECT images were normalized to MNI space (Evans et al., 1994), and average values were calculated for the 163 brain regions of interest.

## ***Gene expression and brain parcellation***

Microarray data was downloaded from the Allen Human Brain Atlas (AHBA) (RRID:[SCR\\_007416](https://www.ebi.ac.uk/RRID/SCR_007416)) website (<http://www.brain-map.org>) ([Hawrylycz et al., 2012](#)). The AHBA data consists of mRNA expression in 3702 tissue samples obtained from six adult human brains, with no known neuropathological history. The data was preprocessed by the Allen Institute to reduce the effects of bias due to batch effects. Description of the processing steps can be found in the technical white paper on AHBA website. For each brain, there are 58,692 probes representing 20,267 unique genes. Leveraging the spatial dependence of gene expression patterns, ([Gryglewski et al., 2018](#)), Gaussian kernel regression was applied to predict the mRNA intensity in each of the 3702 samples in MNI space using leave-one-out cross-validation. The probe with the highest prediction accuracy (among the multiple probes for a gene) was chosen as the representative probe for that gene. Next, because mRNA values were not available for all the grey matter voxels of the brain, Gaussian kernel regression was again used to predict the GE for the remaining MNI coordinates without mRNA expression intensity. Thus, the whole-brain GE data was obtained for the selected 20,267 probes/genes. As it was infeasible to use these ~20,000 AHBA genes for modelling, we therefore selected 976 landmark genes (Table 3.S2) ([Subramanian et al., 2017](#)). These landmark genes are universally informative transcripts with the capacity to cover most of the information in the whole human transcriptome across a diversity of tissue types. The average expression value of each gene was then calculated for the 163 brain regions of interest.

A brain parcellation was derived from a combination of the Jülich, Brodmann, AAL3 and DISTAL atlases. First, structural T1 images of the four atlases were registered to the MNI ICBM152 T1 template using FSL's FLIRT affine registration tool. Then, the obtained transformations were used to project the corresponding parcellations to the MNI ICBM152 space using nearest neighbour interpolation. The resulting parcellation has 163 gray matter

regions of interest which were used to extract the multimodal imaging data, gene expression, and diffusion-based connectivity matrix.

### ***Anatomical connectivity estimation***

The connectivity matrix was constructed in DSI Studio (<http://dsi-studio.labsolver.org>) using a group average template from 1065 subject (Yeh et al., 2018). A multi-shell high-angular-resolution diffusion scheme was used, and the b-values were 990, 1985, and 2980 s/mm<sup>2</sup>. The total number of sampling directions was 270. The in-plane resolution and slice thickness were 1.25 mm. The diffusion data were reconstructed in MNI space using q-space diffeomorphic reconstruction to obtain the spin distribution function (Yeh & Tseng, 2011). The sampling length and output resolution were set to 2.5 and 1 mm, respectively. The restricted diffusion was quantified using restricted diffusion imaging and a deterministic fibre tracking algorithm was used (Yeh et al., 2017). Using the brain atlas previously described under *Methods: Gene Expression and Brain Parcellation*, seeding was placed on the whole brain while setting the QA threshold to 0.15. The angular threshold was randomly varied from 15 to 90 degrees and the step size from 0.5 to 1.5 voxels. The fibre trajectories were smoothed by averaging the propagation direction with a percentage of the previous direction, which was randomly selected from 0 to 95%. Tracks with lengths shorter than 30 mm or longer than 300 mm were discarded. A total of 100,000 tracts were calculated, and the connectivity matrix was obtained by using count of the connecting tracks.

### ***Multimodal neuroimaging modalities***

After preprocessing the imaging modalities, the data were harmonized using ComBat (Fortin et al., 2017). As each site used the same scanner for all subjects, the harmonization procedure corrected for batch effects. The harmonized neuroimaging modalities were extracted for 6 measures, namely dopamine SPECT values, fALLF, fractional anisotropy, mean diffusivity, T1/T2 ratio, and gray matter density. Subjects having at least three neuroimaging

modalities in at least three time points were selected. For these subjects, the modalities missing at each time point having actual individual data were automatically imputed using the trimmed scores regression with internal PCA (Folch-Fortuny et al., 2016). Ultimately, a total of 89 subjects were included in the study with all the 6 neuroimaging modalities for an average of 4 ( $\pm 0.5$ ) time points. The average numbers of imputed time points per neuroimaging modality are presented in Table 3.S1.

### ***Clinical and physical activity measures***

For general clinical measures, we used eleven scores obtained from the PPMI testing battery, namely the Benton Judgment of Line Orientation Test (BJLOT) (Woodard et al., 1996), Hopkins Verbal Learning Test (HVLT) (Brandt, 1991), Letter Number Sequencing (LNS) (Saklofske & Schoenberg, 2011), Geriatric Depression Scale (GDS) (Yesavage, 1988), Movement Disorders Society – Unified Parkinson's Disease Rating Scale (MDS-UPDRS) (Goetz et al., 2008) Parts 1 (non-motor aspects of daily living), 2 (motor aspects of daily living), and 3 (motor examination), the Montreal Cognitive Assessment (MoCA) (Nasreddine et al., 2005), semantic fluency (SF), State-Trait Anxiety Inventory for Adults (STAIAD) (Beckler, 2010), and Symbol Digit Modalities (SDM) (A. Smith, 1973). For the measures of physical activity, we used the three different subscores of the Physical Activity Scale for the Elderly (PASE) (Washburn et al., 1999), with higher scores indicating higher levels of physical activity. The subscores include PASE leisure score, PASE work score, and PASE household score. The methods for deriving all the composite scores are described in the respective PPMI protocols documentation. For each subject, we calculated the rate of change of the scores with respect to the examination date. The slopes of the clinical and physical scores are then used for subsequent analyses.

### 3.6.3 Gene expression multifactorial causal model (GE-MCM)

The GE-MCM models how alterations in different regional neuroimaging-derived biological factors and their interactions are controlled by regional gene expression patterns in the brain (Adewale et al., 2021; Iturria-Medina et al., 2017). Simply, the model is defined by: (i) the influence of each gene on the local direct interactions among all the macroscopic imaging modalities factors, constrained within each brain region, (ii) the potential spreading of macroscopic factor-specific alterations through anatomical and/or vascular networks. (iii) the temporal changes in each macroscopic imaging factor due to (i) and (ii).

In this work, we considered six biological factors namely, brain atrophy, neuronal activity, dopaminergic neuronal loss, dendritic density, and (un)directed measures of white matter integrity. The factors are derived from T1-weighted MR1, resting-state fMR1, DAT-SPECT, T1/T2 ratio, mean diffusivity and fractional anisotropy, respectively. We also considered the regional mRNA patterns of 976 genes. The temporal evolution of the disease-associated process is thus depicted mathematically as:

$$\frac{dS_i^m}{dt} = \sum_{n=1}^{N_{factors}} \left( \alpha_o^{n \rightarrow m} + \sum_{k=1}^{N_{genes}} \alpha_k^{n \rightarrow m} G_i^k \right) S_i^n + \sum_{\substack{j=1 \\ j \neq i}}^{N_{rois}} C_{ji}^m (S_j^m - S_i^m) S^m \quad (1)$$

$N_{genes} = 976$  is the number of genes normalised across  $N_{rois} = 163$  brain regions of interest covering most of the brain's gray matter. Each gene  $i$  is denoted as  $G_i$ , and  $N_{factors} = 6$  is the number of different biological factors measured at the same brain region. The first term on the right-hand side of the equation models the local direct influences of multiple macroscopic biological factors on the given factor  $m$ . The interaction parameters (  $\alpha_o^{n \rightarrow m}$ ,  $\alpha_k^{n \rightarrow m}$  ) and gene expression (  $G_i^k$  ) modulate the direct within-region impact of the factor  $n$  on  $m$ , including intra-factor effects, i.e., when  $n = m$ .  $\sum_{\substack{j=1 \\ j \neq i}}^{N_{rois}} C_{ji}^m (S_j^m - S_i^m) S^m$  reflects the resultant signal propagation of factor  $m$  from region  $i$  to other brain regions through the

physical network  $C_{ji}^m \cdot \frac{dS_i^m}{dt}$  is the local longitudinal alteration of a macroscopic factor  $m$  at region  $i$  due to the foregoing multiscale interactions.

## 3.7 Statistical analyses

### 3.7.1 Model fitting

Using the GE-MCM differential equation, for each subject  $j$  and biological factor  $m$ , we calculated  $\frac{dS_i^m(j)}{dt}$  between each pair of consecutive time points. The regional values obtained were concatenated into a subject-factor-specific vector  $(\frac{dS_i^m(j)}{dt})$  with  $N_{rois} \cdot (N_{times} - 1)$  unique values. This allowed us to formulate the identification of the model parameters  $(\alpha_o^{n \rightarrow m}, \alpha_k^{n \rightarrow m})$  as a regression problem (with  $\frac{dS_i^m(j)}{dt}$  as dependent variable). Due to the high dimensionality of the data, we used a Bayesian sparse linear regression with horseshoe hierarchy to identify the distribution of the model parameters (Carvalho et al., 2010; Makalic & Schmidt, 2015). We then obtained regression coefficients (gene-imaging parameters) as a measure of transcriptomic effect on the interaction of a macroscopic imaging-based factor with the other macroscopic factors, in driving a longitudinal biological factor alteration. We calculated coefficient of determination ( $R^2$ ) for each subject and neuroimaging modality as a measure of model fit.

### 3.7.2 Covariance of gene-neuroimaging parameters with clinical evaluations

Due to high dimensionality, we first reduced the number of gene-imaging parameters by selecting only the stable parameters whose population-wide 95% confidence interval (CI) exclude zero. We then applied singular value decomposition (SVD) to evaluate how the stable transcriptomic-imaging interactions mediate the rates of change in the eleven clinical scores. The aim of SVD is to identify a few pairs of ‘principal components’ that maximize the cross-

correlation between the two sets of variables (i.e., gene-imaging interactions and slope of clinical evaluations). We tested for the significance of the identified principal components (PC) by permuting the mapping the gene-imaging parameters and the clinical scores. The permutation was run 1000 times and principal components with a null  $p < 0.05$  were considered significant. To identify the genes (gene-imaging parameters) with large and reliable contributions on the significant PCs, we drew 1000 bootstrap samples and calculated the bootstrap ratio of the gene-imaging parameters. The bootstrap ratio is obtained by dividing the saliences (contributions) of the gene-imaging parameters by their respective bootstrap standard errors. Top contributing genes were obtained at a bootstrap ratio  $> 1.96$  (corresponding to 95% CI). Diseases associated with the genes were queried from DisGeNET database in Enrichr-KG (Evangelista et al., 2023) at a significance level of  $q\text{-value} < 0.05$ . We derived the PPI networks and the associated WikiPathways from STRING database while setting the PPI confidence score cut-off to 0.7 (Szklarczyk et al., 2021). The hub genes for each PPI network were identified by ranking according to node degrees using cytoHubba plugin (Chin et al., 2014) in Cytoscape (v3.9.1) (Shannon et al., 2003).

### **3.7.3 Covariance of gene-neuroimaging parameters with physical activity**

We again applied SVD to the stable gene-imaging parameters and the slopes of the three different PASE subscores. Significant principal components were obtained by running 1000 permutation iterations and applying a  $p$ -value threshold of 0.05. To identify the top genes mediating physical activity, we drew 1000 bootstrap samples and applied a bootstrap ratio threshold of 1.96 (95% CI). The biological pathways associated with the genes were identified by combining WikiPathways and Reactome databases via the ClueGO (v2.5.9) (Bindea et al., 2009) plugin in Cytoscape. For each of the significant PCs, we evaluated the contribution of each of the PASE subscores by calculating the relative variances along the axis of the PC.

### 3.7.4 Gene perturbation for drug discovery

To discover putative drugs for PD treatment, we sequentially perturbed the gene expressions in both directions. Using Equation (1), gene expression values, most recent neuroimaging measurements, and estimated gene-imaging parameters of each subject, we simulated disease progression for two years, as captured by the longitudinal change of each neuroimaging modality. To perturb a gene, we increased or decreased its expression value by 20% across the population while keeping the values of other genes constant. We then re-simulated the disease progression for 2 years and observed the impact of the perturbation on disease progression. The relative measure of disease progression score is calculated thus:

$$S_{gene} = \frac{\|S_{perturbation}\|_2}{\|S_{normal}\|_2} \quad (2)$$

Where  $S_{gene}$  is the relative disease progression score of a gene in a patient.  $\|S_{perturbation}\|_2$  is the norm of 2-year predictions of the six predicted neuroimaging values obtained by perturbing a gene in each direction.  $\|S_{normal}\|_2$  is the norm of 2-year predictions of all neuroimaging modalities without perturbing any gene.

A gene is therapeutic if  $S_{gene} < 1$ ; i.e., disease progression induced by perturbation is slower compared to actual disease progression. For each of the perturbation direction, we ranked the genes based on the number of subjects where they are predicted to have therapeutic effects. We selected the genes in the top 90<sup>th</sup> percentile and used EnrichR (Chen et al., 2013) to query the Connectivity Map (CMap) (Lamb et al., 2006) drug database. Specifically, the top genes that are therapeutic due to upward and downward perturbation are queried using CMap-Up and CMap-Down databases, respectively. Top putative drugs were then ranked by EnrichR combined score. Using PubChem database (Kim et al., 2023), we retrieved the diseases (Therapeutic Target Database (TTD)) and the pharmacological classifications (Medical Subject Headings (MeSH)) associated with the top drugs.

### 3.8 Supplementary Tables

Table 3.S1: Distribution of number of time points of neuroimaging modalities.

Neuroimaging modality	Average number of time points	Average number of imputed time points
Grey matter density	$3.7 \pm 0.6$	$0.16 \pm 0.37$
fALFF	$2.3 \pm 0.8$	$1.5 \pm 0.76$
DAT-SPECT	$3.3 \pm 0.8$	$0.5 \pm 0.71$
Fractional anisotropy	$2.5 \pm 0.6$	$1.3 \pm 0.51$
Mean diffusivity	$2.5 \pm 0.6$	$1.3 \pm 0.51$
Dendrite density	$2.8 \pm 0.6$	$1.1 \pm 0.25$

Table 3.S2: List of genes used in this study. Please see Excel file

### 3.9 Data and code availability

The three datasets used in this study are available from the PPMI database (neuroimaging and clinical evaluations; <https://www.ppmi-info.org/>), the HCP database (tractography template for connectivity estimation; <http://www.humanconnectomeproject.org/>), and Allen Human Brain Atlas website (gene expression; <http://human.brain-map.org/static/download>). We anticipate that the GE-MCM method will be released soon as part of our available and open-access, user-friendly software (Iturria-Medina et al., 2021) (<https://www.neuropm-lab.com/neuropm-box.html>).

### 3.10 Acknowledgements

The authors thank Dr Konstantinos Poulakis for his advice on statistical analysis. This research was undertaken thanks in part to funding from: the Parkinson Canada and Fonds de recherche du Québec – Santé (FRQS) Graduate Partnership Fellowship awarded to QA, the *Canada First Research Excellence Fund*, awarded to McGill University for the *Healthy Brains for Healthy Lives Initiative*, the Canada Research Chair tier-2, *Fonds de la recherche en santé*

*du Québec* (FRQS) Junior 1 Scholarship, Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant, and Weston Brain Institute awards to YIM, the *Brain Canada Foundation* and *Health Canada* support to the McConnell Brain Imaging Center at the Montreal Neurological Institute, and the *European Union's Horizon 2020 Framework Programme for Research and Innovation* under the Specific Grant Agreements 785907 (Human Brain Project SGA2) and 945539 (Human Brain Project SGA3) awarded to NPG and KZ. Multimodal imaging and clinical data collection and sharing for this project was funded by PPMI. A public-private partnership, PPMI is funded by the Michael J. Fox Foundation for Parkinson's Research and funding partners, including AbbVie, Allergan, Amathus Therapeutics, Avid Radiopharmaceuticals, Biogen, BioLegend, Bristol Myers Squibb, Celgene, Denali Therapeutics, GE Healthcare, Genentech, GlaxoSmithKline plc., Golub Capital, Handl Therapeutics, Insitro, Janssen Neuroscience, Eli Lilly and Company, Lundbeck, Merck Sharp & Dohme Corp., Meso Scale Discovery, Neurocrine Biosciences, Pfizer Inc., Piramal Group, Prevail Therapeutics, Roche, Sanofi Genzyme, Servier Laboratories, Takeda Pharmaceutical Company Limited, Teva Pharmaceutical Industries Ltd., UCB, Verily Life Sciences, and Voyager Therapeutics Inc.

## References

- Adeyale, Q., Khan, A. F., Carbonell, F., & Iturria-Medina, Y. (2021). Integrated transcriptomic and neuroimaging brain model decodes biological mechanisms in aging and Alzheimer's disease. *ELife*, *10*. <https://doi.org/10.7554/eLife.62589>
- Amara, A. W., Chahine, L., Seedorff, N., Caspell-Garcia, C. J., Coffey, C., & Simuni, T. (2019). Self-reported physical activity levels and clinical progression in early Parkinson's disease. *Parkinsonism & Related Disorders*, *61*, 118–125. <https://doi.org/10.1016/J.PARKRELDIS.2018.11.006>
- Armstrong, M. J., & Okun, M. S. (2020). Diagnosis and Treatment of Parkinson Disease: A Review. *JAMA*, *323*(6), 548–560. <https://doi.org/10.1001/JAMA.2019.22360>
- Banerjee, S., Chakraborty, S., Maiti, T., Bisawas, S., & Ganguly College, B. (2021). In silico analysis for potential proteins and microRNAs in Glioblastoma and Parkinsonism.

*BioRxiv*, 2021.10.06.463376. <https://doi.org/10.1101/2021.10.06.463376>

- Beckler, K. (2010). State-Trait Anxiety Inventory for Adults Sampler Set Manual, Instrument and Scoring Guide. *1983 Consulting Psychologists Press, Inc. Mind Garden, Inc.*, 0–78. <https://doi.org/10.1037/T06496-000>
- Bi, M., Du, X., Jiao, Q., Chen, X., & Jiang, H. (2021). Expanding the role of proteasome homeostasis in Parkinson's disease: beyond protein breakdown. *Cell Death & Disease*, *12*(2), 154. <https://doi.org/10.1038/s41419-021-03441-0>
- Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., Fridman, W. H., Pagès, F., Trajanoski, Z., & Galon, J. (2009). ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics*, *25*(8), 1091. <https://doi.org/10.1093/BIOINFORMATICS/BTP101>
- Brandt, J. (1991). The Hopkins Verbal Learning Test: Development of a new memory test with six equivalent forms. *Clinical Neuropsychologist*, *5*(2), 125–142. <https://doi.org/10.1080/13854049108403297>
- Brettle, M., Patel, S., & Fath, T. (2016). Tropomyosins in the healthy and diseased nervous system. *Brain Research Bulletin*, *126*, 311–323. <https://doi.org/10.1016/J.BRAINRESBULL.2016.06.004>
- Butler, E. K., Voigt, A., Lutz, A. K., Toegel, J. P., Gerhardt, E., Karsten, P., Falkenburger, B., Reinartz, A., Winklhofer, K. F., & Schulz, J. B. (2012). The mitochondrial chaperone protein TRAP1 mitigates  $\alpha$ -Synuclein toxicity. *PLoS Genetics*, *8*(2). <https://doi.org/10.1371/JOURNAL.PGEN.1002488>
- Carvalho, C. M., Polson, N. G., & Scott, J. G. (2010). The horseshoe estimator for sparse signals. *Biometrika*, *97*(2), 465–480. <https://doi.org/10.1093/biomet/asq017>
- Chen, E. Y., Tan, C. M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G. V., Clark, N. R., & Ma'ayan, A. (2013). Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics*, *14*, 128. <https://doi.org/10.1186/1471-2105-14-128>
- Chin, C. H., Chen, S. H., Wu, H. H., Ho, C. W., Ko, M. T., & Lin, C. Y. (2014). cytoHubba: Identifying hub objects and sub-networks from complex interactome. *BMC Systems Biology*, *8*(4), 1–7. <https://doi.org/10.1186/1752-0509-8-S4-S11/TABLES/4>
- Collins, G. A., & Goldberg, A. L. (2017). The Logic of the 26S Proteasome. *Cell*, *169*(5), 792–806. <https://doi.org/10.1016/j.cell.2017.04.023>
- Da Silva, P. G. C., Domingues, D. D., De Carvalho, L. A., Allodi, S., & Correa, C. L. (2016). Neurotrophic factors in Parkinson's disease are regulated by exercise: Evidence-based practice. *Journal of the Neurological Sciences*, *363*, 5–15. <https://doi.org/10.1016/J.JNS.2016.02.017>
- Dabool, L., Hakim-Mishnaevski, K., Juravlev, L., Flint-Brodsky, N., Mandel, S., & Kurant, E. (2020). Drosophila Skp1 Homologue SkpA Plays a Neuroprotective Role in Adult Brain. *IScience*, *23*(8), 101375. <https://doi.org/10.1016/j.isci.2020.101375>
- Dong-Chen, X., Yong, C., Yang, X., Chen-Yu, S. T., & Li-Hua, P. (2023). Signaling pathways

- in Parkinson's disease: molecular mechanisms and therapeutic interventions. *Signal Transduction and Targeted Therapy* 2023 8:1, 8(1), 1–18. <https://doi.org/10.1038/s41392-023-01353-3>
- Downs, A. M., Fan, X., Kadakia, R. F., Donsante, Y., Jinnah, H. A., & Hess, E. J. (2021). Cell-intrinsic effects of TorsinA( $\Delta E$ ) disrupt dopamine release in a mouse model of TOR1A dystonia. *Neurobiology of Disease*, 155, 105369. <https://doi.org/https://doi.org/10.1016/j.nbd.2021.105369>
- Dutta, D., Paidi, R. K., Raha, S., Roy, A., Chandra, S., & Pahan, K. (2022). Treadmill exercise reduces  $\alpha$ -synuclein spreading via PPAR $\alpha$ . *Cell Reports*, 40(2). <https://doi.org/10.1016/J.CELREP.2022.111058>
- Elango, R., Banaganapalli, B., Mujalli, A., AlRayes, N., Almaghrabi, S., Almansouri, M., Sahly, A., Jadkarim, G. A., Malik, M. Z., Kutbi, H. I., Shaik, N. A., & Alefishat, E. (2023). Potential Biomarkers for Parkinson Disease from Functional Enrichment and Bioinformatic Analysis of Global Gene Expression Patterns of Blood and Substantia Nigra Tissues. *Bioinformatics and Biology Insights*, 17, 11779322231166214. <https://doi.org/10.1177/11779322231166214>
- Ellis, T., Boudreau, J. K., DeAngelis, T. R., Brown, L. E., Cavanaugh, J. T., Earhart, G. M., Ford, M. P., Foreman, K. B., & Dibble, L. E. (2013). Barriers to Exercise in People With Parkinson Disease. *Physical Therapy*, 93(5), 628. <https://doi.org/10.2522/PTJ.20120279>
- Evangelista, J. E., Xie, Z., Marino, G. B., Nguyen, N., Clarke, D. J. B., & Ma'ayan, A. (2023). Enrichr-KG: bridging enrichment analysis across multiple libraries. *Nucleic Acids Research*, gkad393. <https://doi.org/10.1093/nar/gkad393>
- Evans, A. C., Kamber, M., Collins, D. L., & MacDonald, D. (1994). An MRI-Based Probabilistic Atlas of Neuroanatomy. *Magnetic Resonance Scanning and Epilepsy*, 263–274. [https://doi.org/10.1007/978-1-4615-2546-2\\_48](https://doi.org/10.1007/978-1-4615-2546-2_48)
- Folch-Fortuny, A., Arteaga, F., & Ferrer, A. (2016). Missing Data Imputation Toolbox for MATLAB. *Chemometrics and Intelligent Laboratory Systems*, 154, 93–100. <https://doi.org/10.1016/J.CHEMOLAB.2016.03.019>
- Fortin, J. P., Parker, D., Tunç, B., Watanabe, T., Elliott, M. A., Ruparel, K., Roalf, D. R., Satterthwaite, T. D., Gur, R. C., Gur, R. E., Schultz, R. T., Verma, R., & Shinohara, R. T. (2017). Harmonization of multi-site diffusion tensor imaging data. *NeuroImage*, 161, 149–170. <https://doi.org/10.1016/J.NEUROIMAGE.2017.08.047>
- Frade, J. M., & López-Sánchez, N. (2010). A novel hypothesis for Alzheimer disease based on neuronal tetraploidy induced by p75NTR. *Cell Cycle*, 9(10), 1934–1941. <https://doi.org/10.4161/cc.9.10.11582>
- García-Sanz, P., M.F.G. Aerts, J., & Moratalla, R. (2021). The Role of Cholesterol in  $\alpha$ -Synuclein and Lewy Body Pathology in GBA1 Parkinson's Disease. *Movement Disorders*, 36(5), 1070–1085. <https://doi.org/10.1002/MDS.28396>
- George, G., Valiya Parambath, S., Lokappa, S. B., & Varkey, J. (2019). Construction of Parkinson's disease marker-based weighted protein-protein interaction network for prioritization of co-expressed

genesfile:///C:/Users/elect/Downloads/pmc\_10155030.nbib. *Gene*, 697, 67–77.  
<https://doi.org/https://doi.org/10.1016/j.gene.2019.02.026>

- Ghosh, A., Saminathan, H., Kanthasamy, A., Anantharam, V., Jin, H., Sondarva, G., Harischandra, D. S., Qian, Z., Rana, A., & Kanthasamy, A. G. (2013). The peptidyl-prolyl isomerase Pin1 up-regulation and proapoptotic function in dopaminergic neurons: relevance to the pathogenesis of Parkinson disease. *The Journal of Biological Chemistry*, 288(30), 21955–21971. <https://doi.org/10.1074/jbc.M112.444224>
- Goetz, C. G., Tilley, B. C., Shaftman, S. R., Stebbins, G. T., Fahn, S., Martinez-Martin, P., Poewe, W., Sampaio, C., Stern, M. B., Dodel, R., Dubois, B., Holloway, R., Jankovic, J., Kulisevsky, J., Lang, A. E., Lees, A., Leurgans, S., LeWitt, P. A., Nyenhuis, D., ... Zweig, R. M. (2008). Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results. *Movement Disorders*, 23(15), 2129–2170. <https://doi.org/10.1002/MDS.22340>
- Gryglewski, G., Seiger, R., James, G. M., Godbersen, G. M., Komorowski, A., Unterholzner, J., Michenthaler, P., Hahn, A., Wadsak, W., Mitterhauser, M., Kasper, S., & Lanzenberger, R. (2018). Spatial analysis and high resolution mapping of the human whole-brain transcriptome for integrative analysis in neuroimaging. *NeuroImage*. <https://doi.org/10.1016/j.neuroimage.2018.04.068>
- Hawrylycz, M. J., Lein, E. S., Guillozet-Bongaarts, A. L., Shen, E. H., Ng, L., Miller, J. A., van de Lagemaat, L. N., Smith, K. A., Ebbert, A., Riley, Z. L., Abajian, C., Beckmann, C. F., Bernard, A., Bertagnolli, D., Boe, A. F., Cartagena, P. M., Chakravarty, M. M., Chapin, M., Chong, J., ... Jones, A. R. (2012). An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature*, 489(7416), 391–399. <https://doi.org/10.1038/nature11405>
- He, C. L., Tang, Y., Wu, J. M., Long, T., Yu, L., Teng, J. F., Qiu, W. Q., Pan, R., Yu, C. L., Qin, D. L., Wu, A. G., & Zhou, X. G. (2021). Chlorogenic acid delays the progression of Parkinson's disease via autophagy induction in *Caenorhabditis elegans*. <https://doi.org/10.1080/1028415X.2021.2009993>, 26(1), 11–24. <https://doi.org/10.1080/1028415X.2021.2009993>
- Höglinger, G. U., Adler, C. H., Berg, D., Klein, C., Outeiro, T. F., Poewe, W., Postuma, R., Stoessl, A. J., & Lang, A. E. (2024). A biological classification of Parkinson's disease: the SynNeurGe research diagnostic criteria. *The Lancet Neurology*, 23(2), 191–204. [https://doi.org/https://doi.org/10.1016/S1474-4422\(23\)00404-0](https://doi.org/https://doi.org/10.1016/S1474-4422(23)00404-0)
- Iturria-Medina, Y., Carbonell, F., Assadi, A., Adewale, Q., Khan, A. F., Baumeister, T. R., & Sanchez-Rodriguez, L. (2021). Integrating molecular, histopathological, neuroimaging and clinical neuroscience data with NeuroPM-box. *Communications Biology* 2021 4:1, 4(1), 1–20. <https://doi.org/10.1038/s42003-021-02133-x>
- Iturria-Medina, Y., Carbonell, F. M., Sotero, R. C., Chouinard-Decorte, F., & Evans, A. C. (2017). Multifactorial causal model of brain (dis)organization and therapeutic intervention: Application to Alzheimer's disease. *NeuroImage*. <https://doi.org/10.1016/j.neuroimage.2017.02.058>
- Jin, U., Park, S. J., & Park, S. M. (2019). Cholesterol Metabolism in the Brain and Its

- Association with Parkinson's Disease. *Experimental Neurobiology*, 28(5), 554–567. <https://doi.org/10.5607/EN.2019.28.5.554>
- Khan, A. F., Adewale, Q., Baumeister, T. R., Carbonell, F., Zilles, K., Palomero-Gallagher, N., Iturria-Medina, Y., & Initiative, the A. D. N. (2021). Personalized brain models identify neurotransmitter receptor changes in Alzheimer's disease. *Brain*. <https://doi.org/10.1093/brain/awab375>
- Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., Li, Q., Shoemaker, B. A., Thiessen, P. A., Yu, B., Zaslavsky, L., Zhang, J., & Bolton, E. E. (2023). PubChem 2023 update. *Nucleic Acids Research*, 51(D1), D1373–D1380. <https://doi.org/10.1093/NAR/GKAC956>
- Koo, J. H., Jang, Y. C., Hwang, D. J., Um, H. S., Lee, N. H., Jung, J. H., & Cho, J. Y. (2017). Treadmill exercise produces neuroprotective effects in a murine model of Parkinson's disease by regulating the TLR2/MyD88/NF- $\kappa$ B signaling pathway. *Neuroscience*, 356, 102–113. <https://doi.org/10.1016/J.NEUROSCIENCE.2017.05.016>
- Lamb, J., Crawford, E. D., Peck, D., Modell, J. W., Blat, I. C., Wrobel, M. J., Lerner, J., Brunet, J. P., Subramanian, A., Ross, K. N., Reich, M., Hieronymus, H., Wei, G., Armstrong, S. A., Haggarty, S. J., Clemons, P. A., Wei, R., Carr, S. A., Lander, E. S., & Golub, T. R. (2006). The connectivity map: Using gene-expression signatures to connect small molecules, genes, and disease. *Science*, 313(5795), 1929–1935. [https://doi.org/10.1126/SCIENCE.1132939/SUPPL\\_FILE/LAMB.SOM.PDF](https://doi.org/10.1126/SCIENCE.1132939/SUPPL_FILE/LAMB.SOM.PDF)
- Langeskov-Christensen, M., Franzén, E., Hvid, L. G., & Dalgas, U. (2024). Exercise as medicine in Parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry*, jnnp-2023-332974. <https://doi.org/10.1136/jnnp-2023-332974>
- Li, J., Liu, H., Wang, X., Xia, Y., Huang, J., Wang, T., Lin, Z., & Xiong, N. (2022). Melatonin ameliorates Parkinson's disease via regulating microglia polarization in a ROR $\alpha$ -dependent pathway. *Npj Parkinson's Disease*, 8(1), 90. <https://doi.org/10.1038/s41531-022-00352-5>
- Liou, Y. C., Sun, A., Ryo, A., Zhou, X. Z., Yu, Z. X., Huang, H. K., Uchida, T., Bronson, R., Bing, G., Li, X., Hunter, T., & Lu, K. P. (2003). Role of the prolyl isomerase Pin1 in protecting against age-dependent neurodegeneration. *Nature* 2003 424:6948, 424(6948), 556–561. <https://doi.org/10.1038/nature01832>
- Liu, X., Chen, J., Guan, T., Yao, H., Zhang, W., Guan, Z., & Wang, Y. (2019). miRNAs and target genes in the blood as biomarkers for the early diagnosis of Parkinson's disease. *BMC Systems Biology*, 13(1), 10. <https://doi.org/10.1186/s12918-019-0680-4>
- Mak, M. K., Wong-Yu, I. S., Shen, X., & Chung, C. L. (2017). Long-term effects of exercise and physical therapy in people with Parkinson disease. *Nature Reviews Neurology* 2017 13:11, 13(11), 689–703. <https://doi.org/10.1038/nrneurol.2017.128>
- Makalic, E., & Schmidt, D. F. (2015). *A simple sampler for the horseshoe estimator*.
- Mestre, T. A., Fereshtehnejad, S. M., Berg, D., Bohnen, N. I., Dujardin, K., Erro, R., Espay, A. J., Halliday, G., Van Hilten, J. J., Hu, M. T., Jeon, B., Klein, C., Leentjens, A. F. G., Marinus, J., Mollenhauer, B., Postuma, R., Rajalingam, R., Rodríguez-Violante, M.,

- Simuni, T., ... Marras, C. (2021). Parkinson's Disease Subtypes: Critical Appraisal and Recommendations. *Journal of Parkinson's Disease*, 11(2), 395–404. <https://doi.org/10.3233/JPD-202472>
- Musanti, R., Parati, E., Lamperti, E., & Ghiselli, G. (1993). Decreased cholesterol biosynthesis in fibroblasts from patients with Parkinson disease. *Biochemical Medicine and Metabolic Biology*, 49(2), 133–142. <https://doi.org/10.1006/BMMB.1993.1016>
- Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53(4), 695–699. <https://doi.org/10.1111/J.1532-5415.2005.53221.X>
- Patel, J. R., McCandless, E. E., Dorsey, D., & Klein, R. S. (2010). CXCR4 promotes differentiation of oligodendrocyte progenitors and remyelination. *Proceedings of the National Academy of Sciences of the United States of America*, 107(24), 11062–11067. <https://doi.org/10.1073/pnas.1006301107>
- Paul, K. C., Chuang, Y. H., Shih, I. F., Keener, A., Bordelon, Y., Bronstein, J. M., & Ritz, B. (2019). The association between lifestyle factors and Parkinson's disease progression and mortality. *Movement Disorders*, 34(1), 58–66. <https://doi.org/10.1002/MDS.27577>
- Ping, Z., Xiaomu, W., Xufang, X., & Liang, S. (2019). Vinpocetine regulates levels of circulating TLRs in Parkinson's disease patients. *Neurological Sciences : Official Journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology*, 40(1), 113–120. <https://doi.org/10.1007/s10072-018-3592-y>
- Pinto, A., Cangelosi, A., Geoghegan, P. A., & Goldstein, J. (2017). Dexamethasone prevents motor deficits and neurovascular damage produced by shiga toxin 2 and lipopolysaccharide in the mouse striatum. *Neuroscience*, 344, 25–38. <https://doi.org/10.1016/j.neuroscience.2016.12.036>
- Righart, R., Biberacher, V., Jonkman, L. E., Klaver, R., Schmidt, P., Buck, D., Berthele, A., Kirschke, J. S., Zimmer, C., Hemmer, B., Geurts, J. J. G., & Mühlau, M. (2017). Cortical pathology in multiple sclerosis detected by the T1/T2-weighted ratio from routine magnetic resonance imaging. *Annals of Neurology*, 82(4), 519–529. <https://doi.org/10.1002/ana.25020>
- Saklofske, D. H., & Schoenberg, M. R. (2011). Wechsler Adult Intelligence Scale (All Versions). *Encyclopedia of Clinical Neuropsychology*, 2675–2680. [https://doi.org/10.1007/978-0-387-79948-3\\_1073](https://doi.org/10.1007/978-0-387-79948-3_1073)
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., & Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research*, 13(11), 2498–2504. <https://doi.org/10.1101/GR.1239303>
- Sharma, N., Soni, R., Sharma, M., Chatterjee, S., Parihar, N., Mukarram, M., kale, R., Sayyed, A. A., Behera, S. K., & Khairnar, A. (2022). Chlorogenic Acid: a Polyphenol from Coffee Rendered Neuroprotection Against Rotenone-Induced Parkinson's Disease by GLP-1 Secretion. *Molecular Neurobiology*, 59(11), 6834–6856. <https://doi.org/10.1007/S12035->

- Sharma, R., Kumar, D., Jha, N. K., Jha, S. K., Ambasta, R. K., & Kumar, P. (2017). Re-expression of cell cycle markers in aged neurons and muscles: Whether cells should divide or die? *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, *1863*(1), 324–336. <https://doi.org/https://doi.org/10.1016/j.bbadis.2016.09.010>
- Simon, D. K., Tanner, C. M., & Brundin, P. (2020). Parkinson Disease Epidemiology, Pathology, Genetics and Pathophysiology. *Clinics in Geriatric Medicine*, *36*(1), 1. <https://doi.org/10.1016/J.CGER.2019.08.002>
- Simuni, T., Chahine, L. M., Poston, K., Brumm, M., Buracchio, T., Campbell, M., Chowdhury, S., Coffey, C., Concha-Marambio, L., Dam, T., DiBiaso, P., Foroud, T., Frasier, M., Gochanour, C., Jennings, D., Kiebertz, K., Kopil, C. M., Merchant, K., Mollenhauer, B., ... Marek, K. (2024). A biological definition of neuronal  $\alpha$ -synuclein disease: towards an integrated staging system for research. *The Lancet Neurology*, *23*(2), 178–190. [https://doi.org/https://doi.org/10.1016/S1474-4422\(23\)00405-2](https://doi.org/https://doi.org/10.1016/S1474-4422(23)00405-2)
- Singh, S. S., Rai, S. N., Birla, H., Zahra, W., Kumar, G., Gedda, M. R., Tiwari, N., Patnaik, R., Singh, R. K., & Singh, S. P. (2018). Effect of chlorogenic acid supplementation in MPTP-intoxicated mouse. *Frontiers in Pharmacology*, *9*(AUG), 757. <https://doi.org/10.3389/FPHAR.2018.00757/BIBTEX>
- Smith, A. (1973). *Symbol digit modalities test*. Western psychological services Los Angeles.
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., Bannister, P. R., De Luca, M., Drobnjak, I., Flitney, D. E., Niazy, R. K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J. M., & Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, *23*, S208–S219. <https://doi.org/https://doi.org/10.1016/j.neuroimage.2004.07.051>
- Srinivasan, V., Cardinali, D. P., Srinivasan, U. S., Kaur, C., Brown, G. M., Spence, D. W., Hardeland, R., & Pandi-Perumal, S. R. (2011). Therapeutic potential of melatonin and its analogs in Parkinson's disease: focus on sleep and neuroprotection. *Therapeutic Advances in Neurological Disorders*, *4*(5), 297–317. <https://doi.org/10.1177/1756285611406166>
- Subramanian, A., Narayan, R., Corsello, S. M., Peck, D. D., Natoli, T. E., Lu, X., Gould, J., Davis, J. F., Tubelli, A. A., Asiedu, J. K., Lahr, D. L., Hirschman, J. E., Liu, Z., Donahue, M., Julian, B., Khan, M., Wadden, D., Smith, I. C., Lam, D., ... Golub, T. R. (2017). A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. *Cell*, *171*(6), 1437–1452.e17. <https://doi.org/10.1016/J.CELL.2017.10.049>
- Szklarczyk, D., Gable, A. L., Nastou, K. C., Lyon, D., Kirsch, R., Pyysalo, S., Doncheva, N. T., Legeay, M., Fang, T., Bork, P., Jensen, L. J., & von Mering, C. (2021). The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Research*, *49*(D1), D605–D612. <https://doi.org/10.1093/nar/gkaa1074>
- Tang, L., Fang, Y., & Yin, J. (2019). The effects of exercise interventions on Parkinson's

- disease: A Bayesian network meta-analysis. *Journal of Clinical Neuroscience : Official Journal of the Neurosurgical Society of Australasia*, 70, 47–54. <https://doi.org/10.1016/J.JOCN.2019.08.092>
- Wakabayashi-Ito, N., Doherty, O. M., Moriyama, H., Breakefield, X. O., Gusella, J. F., O'Donnell, J. M., & Ito, N. (2011). Dtorsin, the Drosophila ortholog of the early-onset dystonia TOR1A (DYT1), plays a novel role in dopamine metabolism. *PloS One*, 6(10), e26183. <https://doi.org/10.1371/journal.pone.0026183>
- Wang, H., Chen, Y., Chen, J., Zhang, Z., Lao, W., Li, X., Huang, J., & Wang, T. (2014). Cell cycle regulation of DNA polymerase beta in rotenone-based Parkinson's disease models. *PloS One*, 9(10), e109697. <https://doi.org/10.1371/journal.pone.0109697>
- Wang, M., Li, T., Gao, R., Zhang, Y., & Han, Y. (2023). Identifying the potential genes in alpha synuclein driving ferroptosis of Parkinson's disease. *Scientific Reports*, 13(1), 16893. <https://doi.org/10.1038/s41598-023-44124-4>
- Washburn, R. A., McAuley, E., Katula, J., Mihalko, S. L., & Boileau, R. A. (1999). The physical activity scale for the elderly (PASE): evidence for validity. *Journal of Clinical Epidemiology*, 52(7), 643–651. [https://doi.org/10.1016/S0895-4356\(99\)00049-9](https://doi.org/10.1016/S0895-4356(99)00049-9)
- Woodard, J. L., Benedict, R. H., Roberts, V. J., Goldstein, F. C., Kinner, K. M., Capruso, D. X., & Clark, A. N. (1996). Short-form alternatives to the Judgment of Line Orientation Test. *Journal of Clinical and Experimental Neuropsychology*, 18(6), 898–904. <https://doi.org/10.1080/01688639608408311>
- Yeh, F. C., Liu, L., Hitchens, T. K., & Wu, Y. L. (2017). Mapping immune cell infiltration using restricted diffusion MRI. *Magnetic Resonance in Medicine*, 77(2), 603–612. <https://doi.org/10.1002/MRM.26143>
- Yeh, F. C., Panesar, S., Fernandes, D., Meola, A., Yoshino, M., Fernandez-Miranda, J. C., Vettel, J. M., & Verstynen, T. (2018). Population-averaged atlas of the macroscale human structural connectome and its network topology. *NeuroImage*, 178, 57–68. <https://doi.org/10.1016/J.NEUROIMAGE.2018.05.027>
- Yeh, F. C., & Tseng, W. Y. I. (2011). NTU-90: a high angular resolution brain atlas constructed by q-space diffeomorphic reconstruction. *NeuroImage*, 58(1), 91–99. <https://doi.org/10.1016/J.NEUROIMAGE.2011.06.021>
- Yesavage, J. A. (1988). Geriatric Depression Scale. *Psychopharmacology Bulletin*, 24(4), 709–711.
- Yu, D., Lim, J., Wang, X., Liang, F., & Xiao, G. (2017). Enhanced construction of gene regulatory networks using hub gene information. *BMC Bioinformatics*, 18(1), 186. <https://doi.org/10.1186/s12859-017-1576-1>
- Zhang, J.-H., Ni, S.-Y., Tan, Y.-T., Luo, J., & Wang, S.-C. (2022). A bibliometric analysis of PIN1 and cell death. In *Frontiers in Cell and Developmental Biology* (Vol. 10). <https://www.frontiersin.org/articles/10.3389/fcell.2022.1043725>
- Zou, Q. H., Zhu, C. Z., Yang, Y., Zuo, X. N., Long, X. Y., Cao, Q. J., Wang, Y. F., & Zang, Y. F. (2008). An improved approach to detection of amplitude of low-frequency

fluctuation (ALFF) for resting-state fMRI: fractional ALFF. *Journal of Neuroscience Methods*, 172(1), 137–141. <https://doi.org/10.1016/J.JNEUMETH.2008.04.012>

# Chapter 4: Single-nucleus RNA velocity reveals critical synaptic and cell-cycle dysregulations in neuropathologically confirmed Alzheimer's disease

## 4.1 Preamble

In the previous two chapters, we developed and applied a novel dynamical systems model to study AD and PD. Transcriptomic data used were bulk gene expression data, which can mask cell-specific changes. As transcriptomic changes are one of the cell-autonomous processes that influence selective vulnerability of brain cells to neurodegeneration, studying dynamic transcriptomic changes at single-cell level can provide insight into neuronal and glial cell changes to pathology. A few months before the beginning of this PhD work, a dynamical system model called RNA velocity was developed by researchers in the single-cell transcriptomics field. RNA velocity is being continuously improved upon even at the time of writing this thesis. The model has been applied to study the trajectory of cells in development and cancer. In this chapter, we apply RNA velocity to study neurodegenerative disease for the first time. We uncovered cell-type specific changes in AD, as well as the genes and biological mechanisms driving those changes.

This work has been published as:

Adewale, Quadri, Ahmed F. Khan, David A. Bennett, and Yasser Iturria-Medina. 2024. "Single-Nucleus RNA Velocity Reveals Critical Synaptic and Cell-Cycle Dysregulations in Neuropathologically Confirmed Alzheimer's Disease." *Scientific Reports* 2024 14:1 14 (1): 1–11. <https://doi.org/10.1038/s41598-024-57918-x>.

## 4.2 Abstract

Typical differential single-nucleus gene expression (snRNA-seq) analyses in Alzheimer's disease (AD) provide fixed snapshots of cellular alterations, making the accurate detection of temporal cell changes challenging. To characterize the dynamic cellular and transcriptomic differences in AD neuropathology, we apply the novel concept of RNA velocity to the study of single-nucleus RNA from the cortex of 60 subjects with varied levels of AD pathology. RNA velocity captures the rate of change of gene expression by comparing intronic and exonic sequence counts. We performed differential analyses to find the significant genes driving both cell type-specific RNA velocity and expression differences in AD, extensively compared these two transcriptomic metrics, and clarified their associations with multiple neuropathologic traits. The results were cross-validated in an independent dataset. Comparison of AD pathology-associated RNA velocity with parallel gene expression differences reveals sets of genes and molecular pathways that underlie the dynamic and static regimes of cell type-specific dysregulations underlying the disease. Differential RNA velocity and its linked progressive neuropathology point to significant dysregulations in synaptic organization and cell development across cell types. Notably, most of the genes underlying this synaptic dysregulation showed increased RNA velocity in AD subjects compared to controls. Accelerated cell changes were also observed in the AD subjects, suggesting that the precocious depletion of precursor cell pools might be associated with neurodegeneration. Overall, this study uncovers active molecular drivers of the spatiotemporal alterations in AD and offers novel insights towards gene- and cell-centric therapeutic strategies accounting for dynamic cell perturbations and synaptic disruptions.

### 4.3 Introduction

Recent advances in single-nucleus RNA sequencing (snRNA-seq) have provided an unprecedented ability to disentangle cellular-level transcriptomic alterations and heterogeneity in Alzheimer's disease (AD) (Lau et al., 2020; Mathys et al., 2019). Early differential expression with neuropathological AD progression have been found to be cell-type dependent while apparent upregulation of genes in the later stages are shared across cell types, suggesting that transcriptional responses to disease are highly driven by cell states (Mathys et al., 2019). Furthermore, cell type-specific analysis has revealed the molecular signatures of preferentially affected cell populations, demonstrating that morphology alone cannot sufficiently determine cell type vulnerability in pathologic AD (Bergen et al., 2020; Leng et al., 2021; Olah et al., 2020).

Nevertheless, static snRNA-seq abundance provides only fixed snapshots of cellular states, not revealing temporal dynamics of genes at individual cells (Bergen et al., 2020). The recently proposed rate of change of mRNA, otherwise known as RNA velocity (RNA-vel) (La Manno et al., 2018), provides a novel method to capture temporal dynamics in mRNA abundance by comparing spliced and unspliced mRNA counts. In the initial model, ratio of intronic to exonic sequence counts in constant (steady-state) transcription is obtained, and RNA-vel is estimated as the deviation or residual of this ratio from the expected steady-state ratio (La Manno et al., 2018). Further methods have been developed to capture (potentially) unobserved steady states and gene stochasticity (Bergen et al., 2020). Positive and negative RNA velocities imply upregulation and downregulation of a gene, respectively. Notably, RNA-vel analysis was used to infer developmental trajectories of healthy cells (Kanton et al., 2019; Lo Giudice et al., 2019) and unravel pathological changes in cancer cells (Couturier et al., 2020). This paradigm shift from descriptive to predictive RNA modelling is offering a deeper

understanding of complex cell-level processes in health and disease, with promising implications to improve treatment strategies for multiple disorders.

In the context of the progression of neuropathology, it is unclear whether genes that are differentially expressed across disease states are also the genes driving evolution and vulnerability of diseased cells. Genes with different RNA velocity values might better capture or capture complementary aspects of the time-resolved molecular dysregulations and prodromal differences underlying neurodegeneration. Here, we extend previous single-nucleus RNA (snRNA) analysis in AD in three fundamental ways. First, we use postmortem snRNA-seq data from the prefrontal cortex of subjects with varied levels of AD pathology (N=48) to identify cell type-specific RNA velocity differences associated with neuropathology. Second, we demonstrate that dynamically altered genes in AD pathology, i.e., genes with differential RNA velocities, are qualitatively different from the genes showing differential expression patterns. Third, we reproduce the main observed cell type-specific RNA velocity differences in an independent pathologic AD sample. Overall, our results highlight the critical importance of further considering dynamic single-cell molecular processes underlying AD progression as opposed to only its static cellular RNA mechanisms.

## **4.4 Results**

### **4.4.1 Data origin and single-nucleus RNA velocity estimation**

Single nucleus RNA-seq data was obtained from the prefrontal cortex of 48 postmortem human brain samples (Mathys et al., 2019) (*Methods, Dataset-1* subsection). Twenty-four of these individuals had no or low  $\beta$ -amyloid burden or other pathologies (control). The remaining twenty-four presented mild to severe AD-pathology (amyloid burden, neurofibrillary tangles, global pathology, and cognitive impairment). After pre-processing, 65,422 snRNA-seq profiles with 16,844 transcripts (corresponding to 16,829 unique genes) were obtained. A predefined cluster list (Mathys et al., 2019) was used to annotate and assign the cells to six different types:

excitatory neurons, inhibitory neurons, astrocytes, microglia, oligodendrocytes, and oligodendrocyte progenitor cells (Figure 4.1a).

Single-nucleus RNA velocities (snRNA-vel) of all genes across each cell type in the 48 subjects were calculated using a probabilistic model (Bergen et al., 2020) (*Methods, RNA velocity estimation* subsection). This probabilistic (stochastic) method for RNA-vel estimation is preferred over the originally proposed steady-state model (La Manno et al., 2018) since the former largely accounts for cell heterogeneity and differential kinetics, while achieving higher computational efficiency (Bergen et al., 2020). To identify the genes that may help explain the velocity vector fields across the six types, we selected the top genes that show cell type-specific differential transcriptional dynamics (*Methods, RNA velocity estimation* subsection). As shown in Figure 4.1b, the dependency between unspliced and spliced mRNA counts of the genes gives the expected cell type-specific velocity values depicted as the residual from the dotted line (representing the constant transcriptional state). We then projected the expression and velocity values of these top genes to t-SNE space (Figure 4.1c). We observed more variation and cell type-specificity in RNA velocity compared to gene expression, suggesting that the velocity estimations are largely driven by transcriptional dynamics rather than gene expression (Figure 4.1c). For example, *PDE4B* exhibits an oligodendrocyte-specific dynamics even though its expression is spread across different cell types.

To demonstrate the suitability of using single nuclei for RNA velocity, we first compared both nucleus-derived (snRNA-seq) and cell-derived (scRNA-seq) RNA velocities in the microglia of the same subject. Notably, we observed (Figure 4.S1) strong correlations (ranging from 0.94 to 0.99) between the velocity estimates of the snRNA-seq and scRNA-seq, supporting the precision of snRNA-seq for RNA velocity calculation (see Figure 4.S1). Interestingly, the variations observed in velocity correlations from any random pair of single cells are comparable to the variations in velocity correlations between any single cell and single

nucleus. A previous study also found concordant RNA velocity estimates between matched single nuclei and single cells from rabbit retina (Santiago et al., 2023). Finally, we recalculated RNA velocity using veloVI (Gayoso et al., 2023), another method of RNA velocity estimation based on deep generative modeling. Comparison of the results from scVelo and veloVI shows similar velocity estimates and trajectory inference (Figure 4.S2).

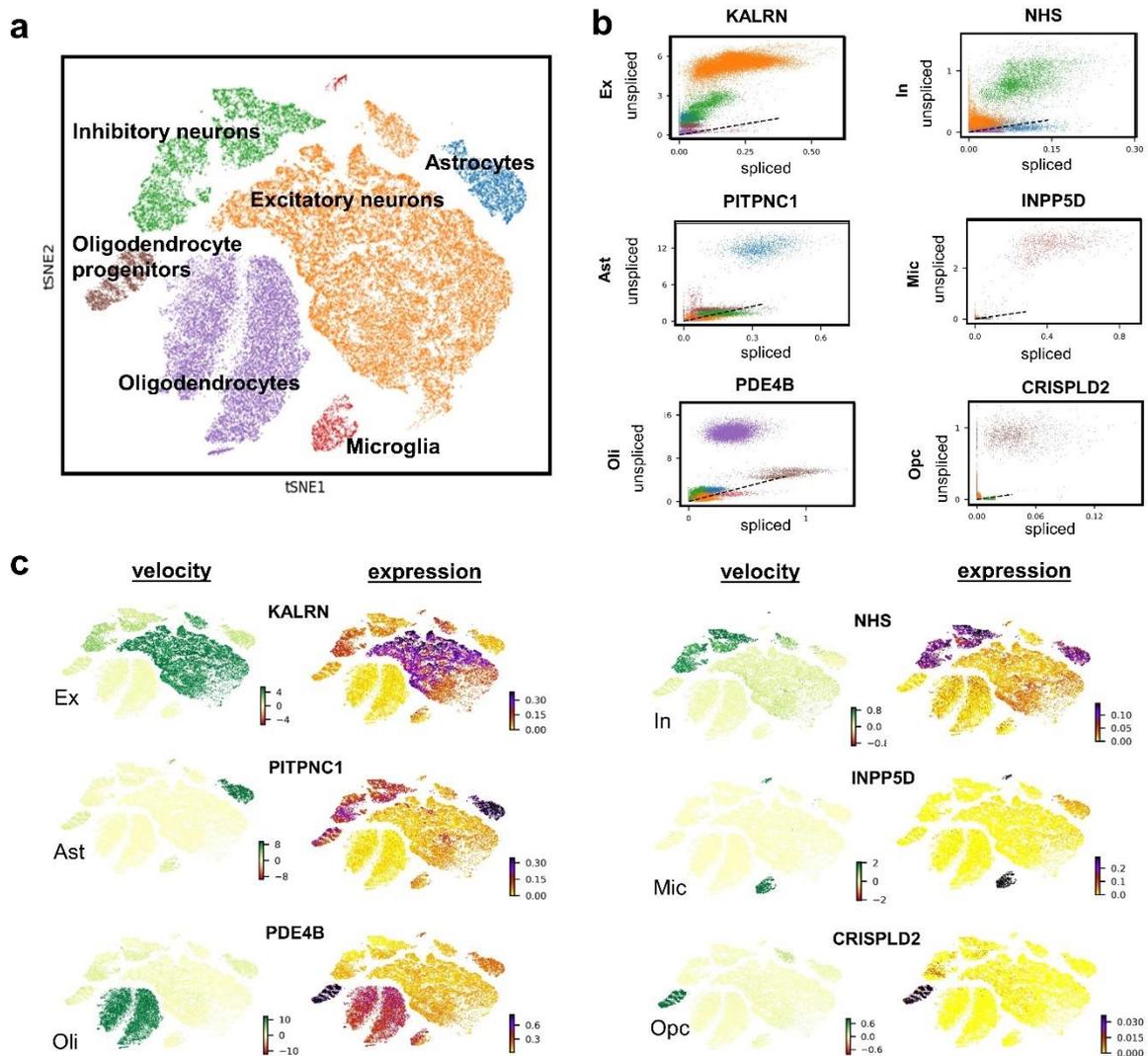


Figure 4.1: Single-nucleus RNA-seq of the prefrontal cortex of 48 individuals across the Alzheimer's disease (AD) spectrum. a) t-SNE visualization of clusters are annotated by cell type (excitatory neurons, inhibitory neurons, astrocytes, microglia, oligodendrocytes, and oligodendrocyte progenitor cells). b) Relationship between unspliced and spliced mRNA counts of genes driving differential dynamics of each cell type. The dashed black line represents the estimated steady state ratio of the unspliced to spliced mRNA. RNA velocity is obtained as the residual of the observed intronic to exonic RNA ratio from this steady state line. c) RNA velocity and expression patterns of the dynamic genes. Larger variation in velocity is driven by transcriptional dynamics.

#### 4.4.2 Static vs dynamic genetic-cell modifications underlying AD evolution

We sought to investigate if there are global AD-pathology dependent differences in RNA velocity patterns across cell types. We evaluated the cell type-specific differences in RNA velocity between the control and AD-pathology subjects using Wilcoxon rank-sum test. We then compared the results with those obtained for differential gene expression. Across all cell types, we observed lesser genes with differential RNA velocity (612) than those with differential expression (3152) (Figure 4.2a, Supplementary Table 4.1). The top ranked genes underlying RNA expression and velocity variations are presented in Figure 4.2b. Furthermore, higher fold changes were observed for RNA velocities compared to gene expression. To exclude the possible confounding impact of age on the observed differential transcriptional kinetics, we reanalyzed the group differences in RNA velocity after correcting for age. The new result was consistent with the original finding without correction, probably because age was matched between groups (Figure 4.S3). Lastly, due to potential over-representation of long unspliced mRNA transcripts in neurons (Gorin & Pachter, 2021), we checked if the differential velocity observed between the two groups may be biased by gene length. We found no correlation ( $R=0.00081$ ;  $p\text{-value}=0.92$ ) between the U-statistic of Wilcoxon rank-sum test and the length of genes in inhibitory neurons (Figure 4.S4).

Notably, 63 of the 3152 (2%) differentially expressed genes were also found to exhibit differential velocity, suggesting substantial AD-pathology related differences between these two RNA descriptors. On the one hand, the genes with only snRNA-vel differences relate to cell developmental and synaptic processes such as morphogenesis, axonal guidance, ion channel activity, synapse organization and cell assembly (Figure 4.2c, Supplementary Table 4.2). The protein-protein interaction network of genes associated with ion channel activity and synapse organization in excitatory neurons is shown in Figure 4.S5. Conversely, genes with only differential expression are majorly associated with mitochondrial activity, ribosomal

processes, and protein sorting. This mismatch between snRNA-vel and RNA abundance dysregulations suggests that analyzing RNA velocity provides relevant complementary information about the multifactorial molecular processes associated with neuropathological AD advance compared to differential expression.

We next investigated if the changes in RNA velocity depend on disease stage. We subgrouped the AD-pathology subjects into early- and late-AD, based on previous study (Mathys et al., 2019). Briefly, the two pathological subgroups were obtained by clustering the subjects on several clinico-pathological features (Mathys et al., 2019). Consequently, early-AD corresponds to some amyloid load with moderate neurofibrillary tangles and cognitive deficit. Late-AD subjects display higher amyloid load and increased neurofibrillary tangles, global pathology, and cognitive deficit. Comparison of the control and early-AD subjects revealed broad-scale changes in transcriptional kinetics between these subgroups (Supplementary Table 3). Corresponding analysis between control and late-AD subjects showed an increased number of affected genes in microglia, astrocytes, and oligodendrocytes, suggesting a progressive immune dysregulation (Figure 4.S6). However, we did not observe any notable shift in cell-type specificity of the velocity differences with disease progression.

Finally, to discern sex-dimorphic differences in transcriptional dynamics, we recomputed differential RNA velocity between the control and AD-pathology subjects while stratifying the data by sex. We found more genes with differential dynamics in females compared to males across cell types (Supplementary Table 4). Interestingly, previous single-cell studies of differential gene expression also observed more dysregulation in female subjects (Belonwu et al., 2022; Mathys et al., 2019), which may account for the higher disease burden in females. Nevertheless, the implicated pathways between both sexes are qualitatively similar, except for some biological processes such as lymphocyte activation and vascular process which are pronounced in the microglia of male (Figure 4.S7).

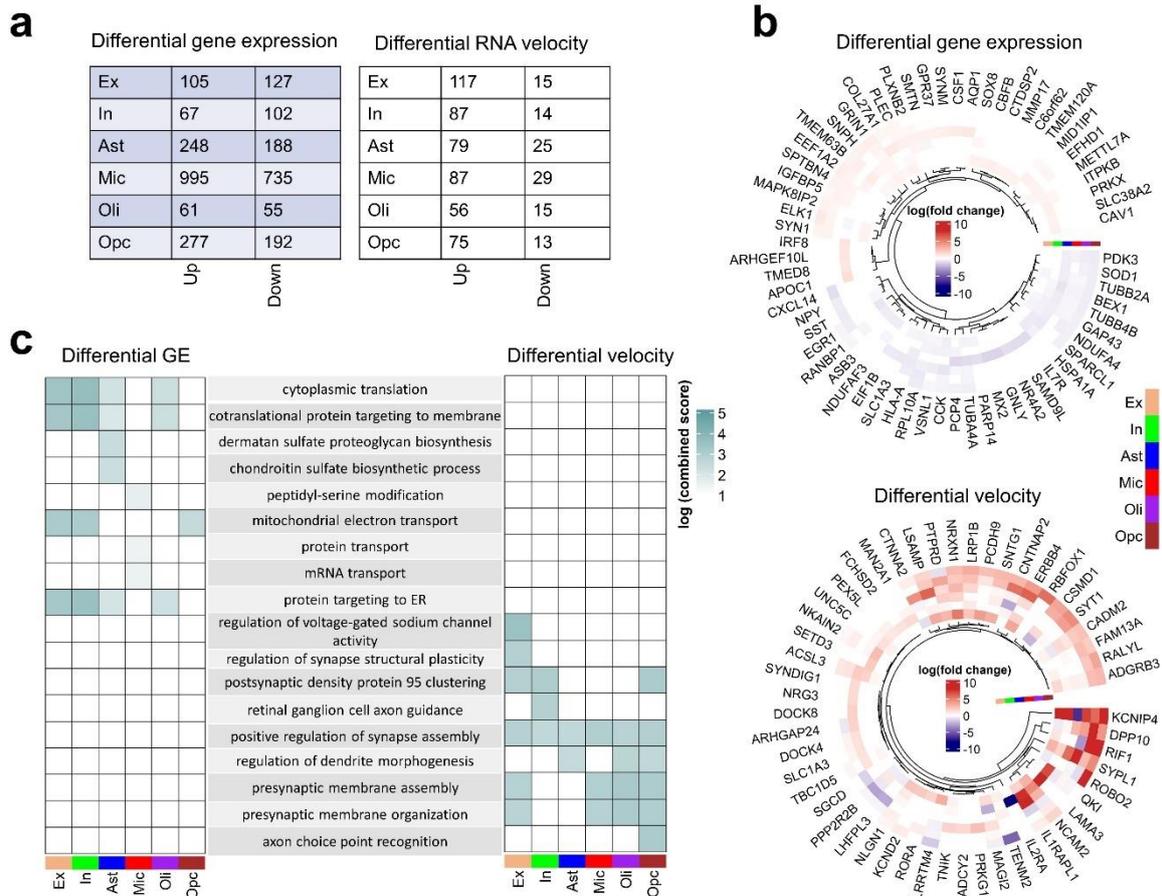


Figure 4.2: Differences in RNA velocity and gene expression underlying neuropathological AD progression. a) Number of genes with differential expression patterns and velocity between controls ( $n=24$ ) and AD subjects ( $n=24$ ) across cell types (two-sided Wilcoxon rank-sum test, permutation-based FDR-corrected  $q$ -value  $< 0.05$ ,  $\log_2(\text{mean gene expression or velocity AD}/\text{mean gene expression or velocity in control}) > 0.25$  or  $< -0.25$ ). b) Cell-type specific changes ( $\log_2(\text{fold change})$ ) for the top genes with differential expression (DE) (top) and differential velocity (bottom) between control and AD subjects. c) Comparison of biological pathways associated with differential expression and differential velocity.

#### 4.4.3 Several RNA-velocity differences underlie AD neuropathological severity

We proceeded to conceptualize the observed differences in RNA velocities in the control and AD groups. We projected the velocity vectors into t-SNE space and evaluated the group difference in velocity fields across all cells (*Methods: Cell speed and residual velocity estimation*). Figure 4.3A shows the residual velocity fields which account for the difference between the two groups. It should be noted that the directions of the residual fields do not have any perceived biological connotation. To further understand the biological implication of the

observed residual velocity fields, we calculated the speed of individual cells using the velocities across all genes. Indeed, we found a higher speed in the AD groups compared to controls, suggesting that accelerated cell changes are associated with AD neuropathology (Figure 4.3b).

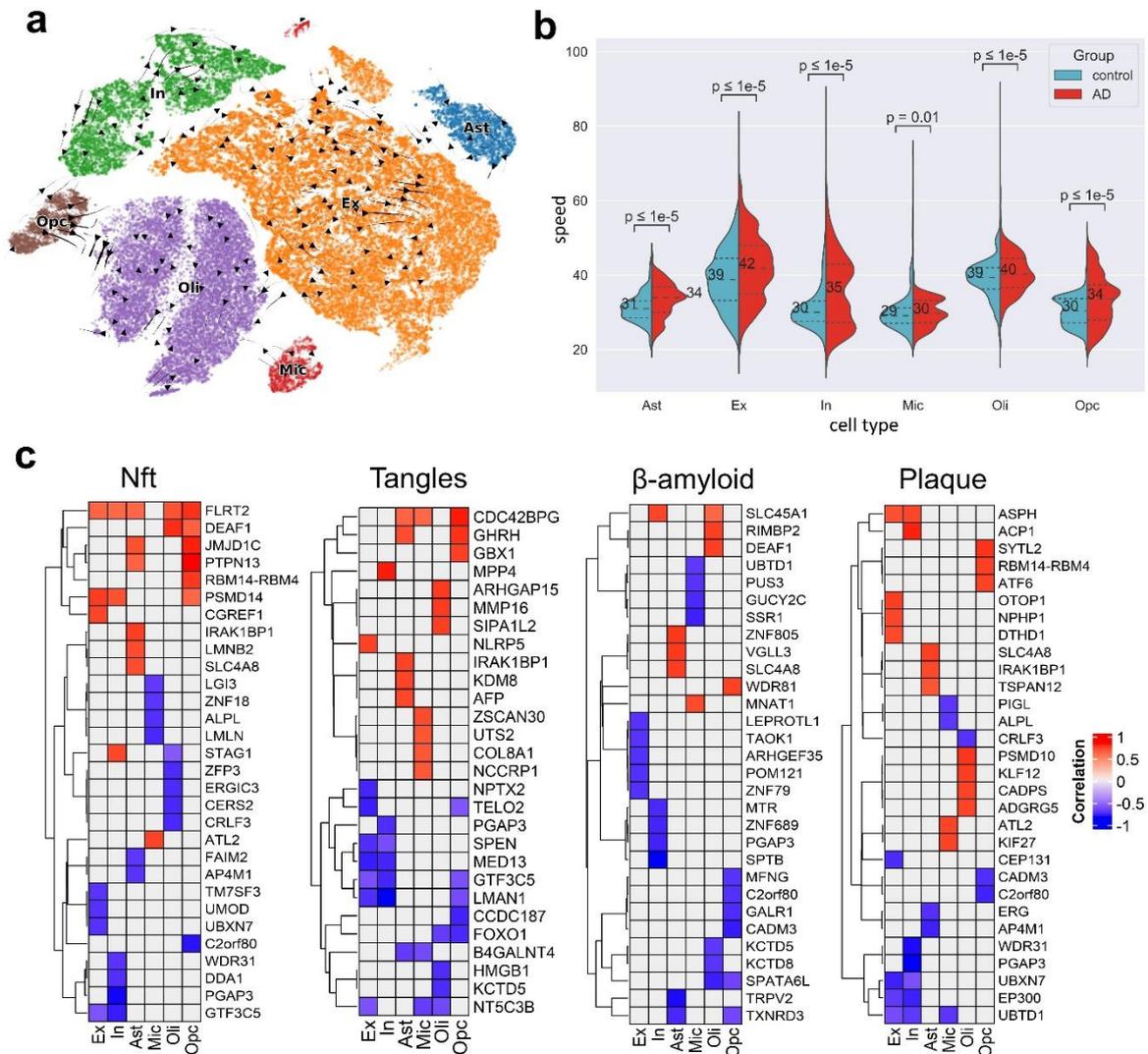


Figure 4.3: Association of RNA velocity with relevant Alzheimer's-related neuropathological traits. a) Residual velocity fields between the control and AD cells ( $z$ -score  $> 1.96$  or  $< -1.96$ ). a) Residual velocity fields between the control and AD cells ( $z$ -score  $> 1.96$  or  $< -1.96$ ). b) Variation of cell speed between the control and AD groups. Median speed overlays plots as numbers (Wilcoxon rank-sum test  $p$ -value  $< 0.05$ ). c) Spearman's correlation of RNA velocity with four AD-neuropathological traits: neuritic plaque count, neuronal neurofibrillary tangle (NFT) counts, overall  $\beta$ -amyloid load ( $\beta$ -amyloid) and PHF tau tangle density (tangles). The confounding effects of age, sex, education, and postmortem interval were accounted for via partial correlation. The plots show the top significant genes at FWE-corrected  $p$ -value  $< 0.001$ .

Lastly, we investigated the association of the RNA velocity differences with four well-known AD pathological traits: neuritic plaque (NP) and neuronal neurofibrillary tangle (NFT)

counts based on histochemistry silver stain, and overall  $\beta$ -amyloid load ( $\beta$ -amyloid) and PHF tau tangle density (tangles) based on molecularly specific immunohistochemistry. We calculated Spearman's correlation between RNA velocity of genes from the 24 AD subjects and the four neuropathological traits while adjusting for the covariates age, sex, education, and postmortem interval. The velocity-phenotype correlations of the top significant genes (FWE-corrected,  $p$ -value $<0.001$ ) are shown in Figure 4.3c. The genes underlying the different neuropathological traits are largely specific to cell types (Figure 4.S8). Only excitatory and inhibitory neurons presented a relatively high overlap (up to 13%) in significant genes associated with the different AD neuropathological traits. The other cell types substantially differed in gene-specific changes across the phenotypes. Interestingly, we found some AD-relevant genes, including *ADAM10*, associated with tangle burden in excitatory and inhibitory neurons. Some other previously reported AD genes identified include amyloid beta precursor protein binding family members, matrix metalloproteinases, notch, low-density lipoproteins, and protein kinase C's (Supplementary Table 5).

#### **4.4.4 Cross-study validation of differential RNA velocity**

We tested the reproducibility of the observed AD-related differences in RNA velocity in an independent sample. We obtained snRNA-seq data from the dorsolateral prefrontal cortex of another ROSMAP cohort (*Dataset-2; Methods, Validation in independent dataset subsection*) comprising 6 cognitively non-impaired individuals with minimal AD pathology and 6 subjects with both pathologic AD and clinical AD dementia (Cain et al., 2020). Following preprocessing, we derived 79,472 cells corresponding to the same six cell types under consideration, and 16,844 transcripts (corresponding to 16,829 unique genes) like in our analysis for the first dataset. Out of the 232 genes with significant RNA velocity differences between the two groups in Dataset-2, 129 (i.e., 56%) overlapped with the genes obtained from the initial analysis of the prefrontal cortex (Dataset-1), including 14 in excitatory neurons, 17

inhibitory neurons, 22 in astrocytes, 34 in microglia, 18 in oligodendrocytes and 24 in oligodendrocytes precursor cells (Figure 4.4a). The significance of overlap was assessed using Fisher’s exact test (p-value<0.01; odds ratio>1).

We proceeded to query enriched gene ontology terms (GO) of the significant genes common to the two datasets. We used EnrichR tool (Chen et al., 2013; Kuleshov et al., 2016) to uncover the biological processes associated with these overlapping genes. Importantly, we again found that most of the top biological processes across cell types are associated with neural development and synaptic activities (Figure 4.4b). Overall, these findings support the generalizability of the main RNA velocity differences identified, supporting the robustness of this novel technique for the deep molecular characterization and better understanding of AD pathomechanisms.

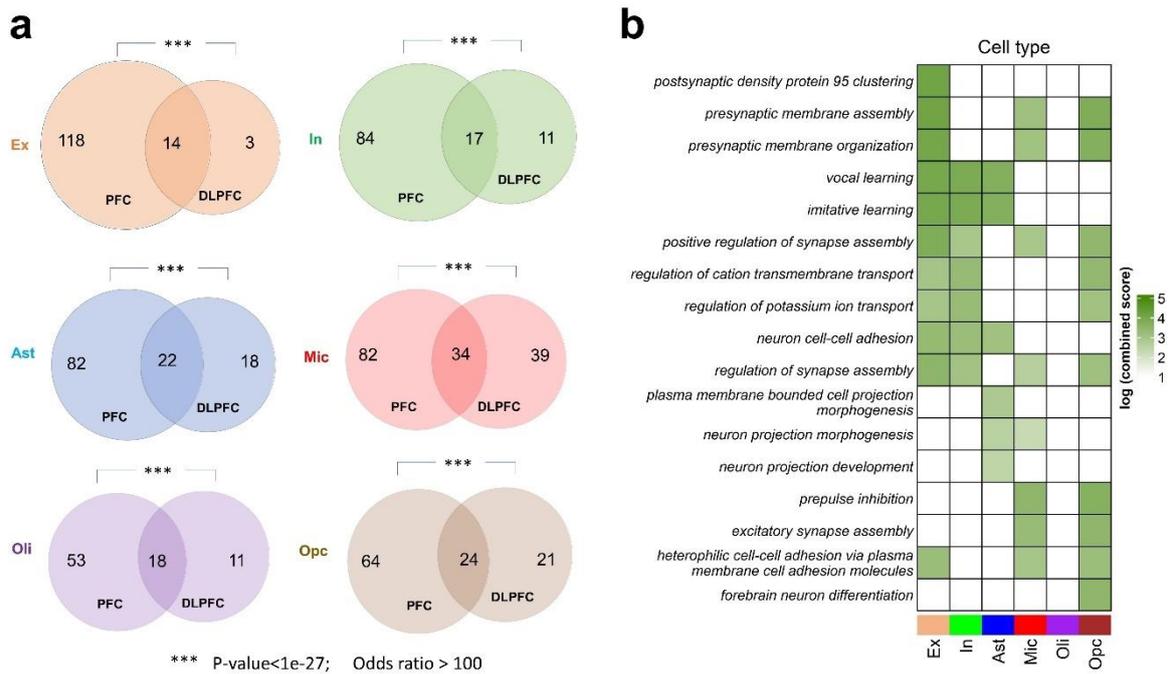


Figure 4.4: Cross-study validation of RNA velocity differences in neuropathologic AD. a) Venn diagrams for each cell type showing the overlaps between Dataset-1 (prefrontal cortex, PFC) and Dataset-2 (dorsolateral prefrontal cortex, DLPC) with respect to genes having differential RNA velocity in AD pathology. The sum of all two numbers in any circle represents the number of significant genes in the corresponding dataset. Significance of overlap was estimated with Fisher’s exact test (p-value<0.01; odds ratio>1). b) GO biological processes and their log-transformed EnrichR combined scores for the overlapping gene sets.

## 4.5 Discussion

Here, we use RNA velocity to characterize, for the first time to our knowledge, the dynamical multicellular processes underlying neuropathological AD progression. Unprecedented advances in scRNA-seq have offered a novel way to overcome the poor spatial resolution of bulk tissue mRNA, while enabling the study of cell type-specific changes in AD and related disorders (Lau et al., 2020; Mathys et al., 2019; Olah et al., 2020). However, differences in RNA expression do not completely capture the evolution of the disease continuum or the progressive vulnerability of cells to neurodegeneration. Using snRNA-seq profiled from postmortem brain samples of the prefrontal cortex and dorsolateral prefrontal cortex in two independent studies, we uncovered highly active genes associated with different levels of neuropathology. Importantly, the identified cross-validated dynamic genes are associated with a consistent set of molecular functions linked with neurodegeneration. The results support the validity of the novel RNA velocity concept for achieving a complementary molecular characterization of AD and potentially identifying cell type-specific disease-modifying genetic targets.

We found accelerated cell dynamics in AD subjects compared to controls, which can explain some of the molecular bases of the early changes occurring in AD. A previous study using induced pluripotent stem cells (iPSC) showed that AD brains undergo accelerated neural differentiation that causes early depletion of neural progenitor pools and reduced cell renewal (Meyer et al., 2019). Further, accelerated cell differentiation may perturb the gene network associated with cell development and synapse organization thereby engendering hyperexcitability and other pathologic cascades. This may as well have implications for the risk of developing dementia as increased cell proliferation of neural progenitor cells in early later and depletion in later life have been linked to *APOE* deficiency (Yang et al., 2011). Our

study suggests that accelerated cell differentiation occurs across different cell types in AD and offers potential areas for experimental validation.

Our analysis revealed that although RNA velocity is closely related to gene expression, the two quantities may capture different pathological processes. The differentially expressed genes associated with AD pathology differed from those associated with varying RNA velocity. The snRNA velocity related genes are principally involved in cell developmental and synaptic programs while the expression-related genes are mainly implicated in ribosomal and mitochondrial activities. However, most of the snRNA synaptic genes are upregulated, except *EPHB1*, *ILIRAPL1* and *ROBO2* in astrocytes and *PTPRD* in excitatory neurons (Supplementary Table 2). Previous studies have shown that astrocytic *EPHB1* and *ROBO2* play vital roles in synapse remodeling and neuronal migration, respectively (Kaneko et al., 2010; Nikolakopoulou et al., 2016). The interaction of *PTPRD* with *ILIRAPL1* promotes excitatory synapse formation and stabilization, and the downregulating either gene impairs synaptogenesis (Blockus et al., 2021; Park et al., 2020; Ramos-Brossier et al., 2015). In addition to  $\beta$ -amyloid and tau related processes, our analysis of snRNA-vel pointed at other potentially altered functions such as voltage-gated cation channels activity and notch signaling which may have dynamic causal roles in AD development but were not detected by the traditional differential expression. Further, we cross-validated the differential RNA velocity analysis in an independent dataset. The overlapping genes between the two datasets are predominantly implicated in biological processes associated with cell development and synapse organization, implying a recurring theme in our results. A profound nexus exists between cell cycle and synaptic activity in AD, and many AD-associated genes are involved in morphoregulation, i.e., the ordered development and arrangement of cells to form synapse through processes such as cell adhesion, cell differentiation, synaptic membrane assembly, ion channel activity, etc. Besides, results from animal studies showed that certain behaviors

simultaneously enhance synaptic plasticity and control accelerated cell cycle, thereby protecting against cell death and neurodegeneration (Arendt, 2003).

The RNA velocity metric was designed to capture the dynamic process of cell evolution in the transcriptomic space (Bergen et al., 2020; La Manno et al., 2018). It was originally applied to infer the developmental states of healthy cells but has found further applications in studying cell proliferation in cancer (Couturier et al., 2020; Pan et al., 2020). However, it appears that RNA velocity can also capture dynamic differences associated with severity of AD pathology. Moreover, RNA velocity can be estimated for each cell type at the patient level. Such applications are particularly important for two main reasons. First, prodromal cell changes which occur in AD may be detected before clinical manifestations or the deposition of  $\beta$ -amyloid and tau (De Strooper & Karran, 2016; Maruszak & Zekanowski, 2011). Second, there are implications for the development of personalized treatment by detecting (and potentially targeting) person-specific contribution of RNA velocity changes to AD neuropathology. We found that most of the biological processes implicated in our study are involved in synapse organization and turnover, a key structural element essential for cognition. Many pathways are also associated with cell developmental processes, the dysfunction of which is linked to neurodegeneration (Joseph et al., 2020). Thus, our results inform potential therapeutic strategies of targeting substrates of synaptic plasticity, including glutamatergic and cholinergic signaling, and applying cell therapy to enhance cell renewal, differentiation, and proliferation.

This study has some limitations. First, we used single-nucleus RNA sequencing to estimate RNA velocity. Compared to single-cell sequencing, snRNA-seq is more amenable to transcriptomic profiling of postmortem samples because isolated nuclei are intact in frozen tissues (Lake et al., 2017). Moreover, dissociating whole cells from the brain is particularly challenging due to the intensity of the required enzymatic activity, which could interfere with

cell type recovery and bias the results of downstream analyses (Habib et al., 2017). RNA velocity was originally formulated for scRNA-seq based on the assumption that the rate of RNA degradation is constant across all cells (La Manno et al., 2018). Importantly, here we showed that there is high concordance between RNA velocities from matched nuclear and whole-cell RNA of microglia, supporting the validity of using snRNA-seq for velocity estimation (see Figure 4.S1). Furthermore, by using two different RNA velocity methods to achieve a consistent trajectory inference, we demonstrated the robustness of the velocities estimated from single-nucleus data. In line with our observations, a recent study comparing RNA velocity trajectory inferences from scRNA-seq and snRNA-seq showed that each method successfully predicts the transition of retinal resting cells through reactive state to terminal fibrous state. Interestingly, the RNA velocity estimates obtained by combining the two sequencing technologies were consistent with the estimates derived from using either scRNA-seq or snRNA-seq dataset alone. Earlier studies have also successfully applied RNA velocity to single nuclei data to infer biologically meaningful trajectories in mouse embryo and various cell types of the human heart (Marsh & Blelloch, 2020; Wolfien et al., 2020). Second, RNA velocity was originally developed to capture rapidly evolving processes with short timescales. Its use in slowly evolving processes, such as neurodegeneration, remains to be validated. However, we believe that the relatively shorter timescale of mRNA transcriptional dynamics may offer a good resolution to capture subtle changes associated with the neuropathological cascade. Despite that the use many of subjects with varied levels of neuropathology allowed us to capture the association between the timescales of transcriptional dynamics and neurodegeneration at the global level, due to experimental limitations we could not directly ascertain the stability of RNA velocities within a subject over different post-mortem intervals. However, we confirmed that the post-mortem sampling intervals between the controls and AD subjects do not differ significantly (Figure 4.S9), suggesting that our results were not

confounded by transcriptional changes induced by death. Importantly, future studies employing multiscale dynamical models of the brain (via gene expression, neuroreceptors, neuroimaging) can also incorporate RNA-vel to better capture the time-resolved complex interactions between multiple biological layers/modalities in neurodegenerative progression (Adewale et al., 2021; Khan et al., 2021).

## 4.6 Methods

### 4.6.1 Dataset-1 (Prefrontal cortex)

It includes droplet-based snRNA-seq, neuropathological and clinical data for 48 participants enrolled in the Religious Orders Study (ROS) or the Rush Memory and Aging Project Study (MAP) (Bennett et al., 2018). The snRNA-seq data was previously generated from the prefrontal cortex (Brodmann area 10) of autopsied brains as described in (Mathys et al., 2019), and it was downloaded from the Accelerating Medicines Partnership Alzheimer's Disease knowledge portal (AMP-AD; [www.synapse.org](http://www.synapse.org), ID syn18485175). All subjects underwent postmortem neuropathologic evaluations, generated in previous ROSMAP studies as described in (Bennett et al., 2018) including uniform structured assessment of AD pathology, and other pathologies common in aging and dementia (downloaded from AMP-AD, ID syn3157322; see also *Correlation with neuropathology* subsection below). The 48 subjects (balanced between sexes) comprised 24 with no or low AD-pathology (control group), and 24 with mild to severe AD-pathology (AD group) as determined by  $\beta$ -amyloid burden, neurofibrillary tangles, and cognitive impairment (Mathys et al., 2019). The subjects were matched for age (medians 87.1 [no pathologic AD, N=24] and 86.7 [pathologic AD, N=24]) and years of education (medians of 18 [no pathologic AD] and 19.5 [pathologic AD]). Informed consent was obtained from each participant, and the Religious Orders Study and Rush Memory and Aging Project were approved by an Institutional Review Board (IRB) of Rush University Medical Center. The project was performed in accordance with the Declaration of Helsinki. In

addition, subjects signed a repository consent allowing their data to be shared (related documents and requests for data can be obtained at <https://www.radc.rush.edu>).

The process of isolating the nuclei from the postmortem brain tissues was previously detailed (Mathys et al., 2019). Briefly, the brain tissue was homogenized at a very low temperature and incubated. The tissue was then filtered and purified with working solutions. The nuclei were separated through spinning at high speed and counted. The sequencing libraries were constructed with the Chromium Single Cell 3' Reagent Kits v.2 (10x Genomics) and sequenced with the NextSeq 500/550 High Output v2 kits (150 cycles).

#### **4.6.2 RNA abundance and cell type identification**

Intronic and exonic counts were obtained using kb-python (v0.26.3), a wrapper for kallisto and bustools (Bray et al., 2016; Melsted et al., 2021). First, index file of the human genome was generated from the Ensembl human primary reference genome sequence and gene annotation (GRCh38). Then, spliced and unspliced RNA counts were obtained by filtering barcodes with low UMI counts and mapping reads to the index file. The counting process was performed by sequentially running 'kb ref', and 'kb count' (with filter flag set) commands.

Next, we acquired a previous quality controlled list of genes, cells and cell types (Mathys et al., 2019). We then looked for shared genes and cells between our filtered counts and the previously reported list. Thus, we had 65,422 cells with 16,844 transcripts (corresponding to 16,829 unique genes). These cells were then assigned to cell types, based on the reported list, as excitatory neurons, inhibitory neurons, astrocytes, microglia, oligodendrocytes, oligodendrocyte precursor cells, endothelial cells, and pericytes. Endothelial and pericyte cells were subsequently excluded because of their very low counts or absence in some subjects.

### 4.6.3 RNA velocity estimation

We used scVelo (v0.2.3) (Bergen et al., 2020) to calculate RNA velocity. First, the cells were pulled together across all subjects, and each cell was normalized by its total size. The normalization was applied to both spliced and unspliced counts. To estimate RNA velocity using the stochastic method, we computed the means and variances of nearest neighbors of cells in principal component analysis (PCA) space. Here, 100 nearest neighbors and 30 principal components were used. Normalization and moments calculation were achieved through ‘pp.normalize\_per\_cell’ and ‘pp.moments’ commands, respectively. The RNA velocity is then estimated with ‘tl.velocity’ command (setting the mode to ‘stochastic’).

We next sought to validate the estimated RNA velocities by examining the velocity values of the genes driving cell type-specific dynamics. We ran a differential velocity Welch’s t-test with the module ‘scv.tl.rank\_velocity\_genes’ and obtained the top genes (based on t-value) having cell type-specific differential velocity. We then projected the velocities and expression values of the dynamic genes into t-SNE space to examine their variations across cell types.

### 4.6.4 Differential expression and RNA velocity analyses

Cell type-specific gene analysis was performed with Seurat (v4.0.2) and Presto packages in R. We performed differential expression analysis between the control group and the AD group. Each cell was first normalized by its total count over all genes, scaled by 10,000 and log-transformed. Using the ‘FoldChange’ command in Seurat, we performed Wilcoxon rank-sum test to identify differentially expressed genes at  $\log_2(\text{fold change}) > 0.25$  or  $< -0.25$ . We then used the Presto package (due to its speed) to run 5000 random permutations by randomly reassigning the subjects to either the control or AD group. The U-statistics from the permutations were used to generate null distributions and significance p-values. We identified significant genes after adjusting for multiple testing ( $q < 0.05$ , FDR-corrected). To compare

differential expression with differential velocity, the procedure was repeated on RNA velocities to identify the dynamic genes driving the velocity difference between the control and AD groups.

#### **4.6.5 Cell speed and residual velocity estimation**

First, the speed of a cell was calculated as the length of its velocity vector. Wilcoxon ranked-sum test was then used to compare the speed between the two groups. Next, the RNA velocities of individual cells were used to derive velocity fields in t-SNE space. For each of the groups (control and AD), we linearly interpolated the velocity fields at the t-SNE coordinates where actual cells are missing to ensure equal number of velocity fields for each group. We then subtracted the velocity fields at same pair of coordinates for the two groups and obtained the z-scores of the norms of these differences. The velocity field difference of those coordinates where the  $z\text{-score} > 1.96$  or  $< -1.96$  are displayed as the residuals between the two groups.

#### **4.6.6 Correlation with neuropathology**

In each cell type and subject, the average RNA velocity across cells was calculated for every gene. The velocities were tested for correlation with four AD neuropathological traits (Bennett et al., 2018): PHF neurofibrillary tangle density (tangles), neuronal neurofibrillary tangle counts (NFT), overall  $\beta$ -amyloid load ( $\beta$ -amyloid), and neuritic plaque counts (NP). The correlations were adjusted for age, sex, and postmortem interval. Significant genes were chosen based on FWER-corrected p-value  $< 0.001$  (Genovese et al., 2002).

#### **4.6.7 Validation in independent dataset (Dataset-2: Dorsolateral prefrontal cortex)**

The droplet-based snRNA-seq data was previously generated from the dorsolateral prefrontal cortex of autopsied brains as described in (Cain et al., 2020), and it was downloaded from the Accelerating Medicines Partnership Alzheimer's Disease knowledge portal (AMP-AD; [www.synapse.org](http://www.synapse.org), ID syn16780177). The subjects include another 12 sex-matched

individuals from Orders Study (ROS) or the Rush Memory and Aging Project Study (MAP) (Bennett et al., 2018): 6 subjects are cognitively non-impaired with minimal AD pathology and 6 fulfil diagnoses for both pathologic AD and clinical AD dementia.

As previously described (Cain et al., 2020), the brain grey matter tissue was homogenized and treated with working solution to separate the nuclei. The isolated nuclei were then counted and filtered. The libraries were constructed and sequenced on the 10X Single Cell RNA-Seq Platform using the Chromium Single Cell 3' Reagent Kits v2. After obtaining the intronic and exonic counts, genes were selected according to the gene list from Dataset-1 and the cells were filtered using the cell list obtained from the metadata of the previous study (Cain et al., 2020). The previously reported cell clusters were used to assign the 79,472 cells to the six cell types under consideration. We calculated the cell type-specific RNA velocities for each subject and used Wilcoxon rank-sum test to identify the genes underlying RNA velocity differences between the minimal AD pathology group and the pathologic/AD dementia group.

We assessed the overlap between the significant differential velocity genes in this dataset and Dataset-1. Fisher's exact test was used to obtain the significance of overlap ( $p\text{-value} < 0.01$  and  $\text{odds ratio} > 1$ ).

#### **4.6.8 Biological pathway analyses**

Biological pathways were identified using EnrichR online tool to query enriched gene ontology (GO) terms (Chen et al., 2013; Kuleshov et al., 2016) from the Gene Ontological Biological Processes 2021. The significant GO terms were selected at an adjusted  $p\text{-value} < 0.01$  and ranked based on their EnrichR combined scores.

#### **4.6.9 Comparison between single-cell and single-nucleus RNA velocities (Dataset-3)**

We downloaded a previously published dataset from the Gene Expression Omnibus (GEO) (<https://www.ncbi.nlm.nih.gov/geo/>, GSE135618). The dataset contains matched single cells, fresh nuclei and frozen nuclei obtained from the microglia of two subjects (Gerrits et al., 2020). After preprocessing, we obtained 2,988 cells, 4,892 fresh nuclei, and 4,019 frozen nuclei from one subject; and 3,485 cells, 2,593 fresh nuclei, and 5,527 frozen nuclei from the other subject. For each of the subjects, we estimated the RNA velocities across the three modalities (single cell, fresh nuclei, and frozen nuclei) and performed within-subject comparisons of the velocity estimates.

#### **4.6.10 Data visualization**

Visualizations were performed using scVelo (v.0.2.3) (Bergen et al., 2020), ComplexHeatMap (v.2.6.2) (Zuguang et al., 2016), g:Profiler (Raudvere et al., 2019), Cytoscape App (3.9.1) (Shannon et al., 2003), and STRING (v.12.0) (Szklarczyk et al., 2021).

## **4.7 Supplementary Figures**

## Validation of the correspondence between RNA velocities of scRNA-seq and snRNA-seq in microglia

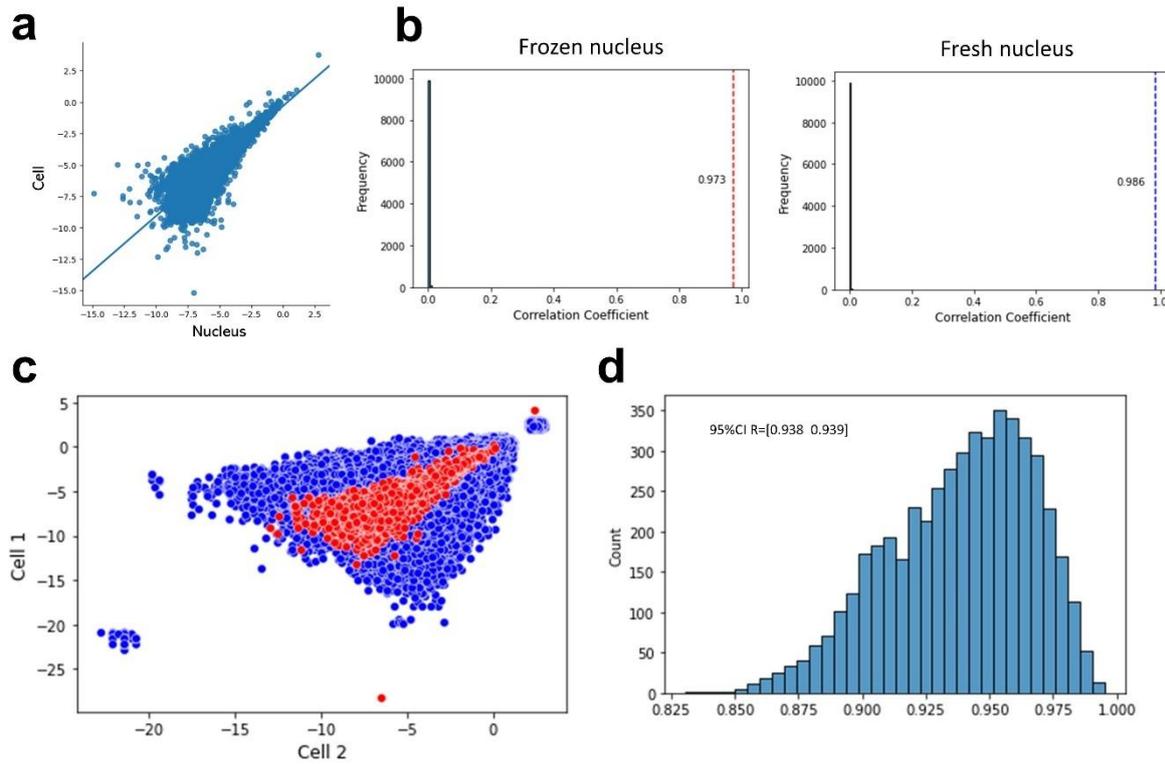


Figure 4.S1: Validation of the correspondence between RNA velocities of single-cell and single-nucleus sequencing dataset from microglia of same subjects. Scatter plot (in log scale) showing the linear concordance between velocities calculated in cells and nuclei for a subject. Each dot represents an independent gene. b) Left: null distribution of Pearson's correlation coefficient between scRNA and frozen snRNA velocities from 5000 permutations. The actual correlation (0.973) is shown in red dotted line. Right: null distribution of Pearson's correlation coefficient between scRNA and fresh snRNA velocities from 5000 permutations. The actual correlation (0.986) is shown in blue dotted line. The second subject (not shown) had correlations of 0.936 and 0.944 in frozen and fresh nuclei, respectively. The high and significant correlations suggest that nuclei RNA can be used to estimate RNA velocities in place of whole-cell RNA. c) Overlay of cell-cell (in blue) and cell-nucleus (in red) RNA velocity correlations. First, a pair of cells is selected at random and the velocities of the genes in the cells are plotted against each other. The random selection is repeated 5000 times to show the variabilities in RNA velocity correlations between single cells. The initial cell-nucleus correlation from A) is then overlayed on the cell-cell correlation plot. The variabilities in cell-nucleus correlation of RNA velocity are observed present in correlations between single cells. D) The correlation distribution between the RNA velocities of any random pair of whole cells. The 95% confidence interval is shown in the plot.

## Comparison of RNA velocity trajectory inference between scVelo and veloVI

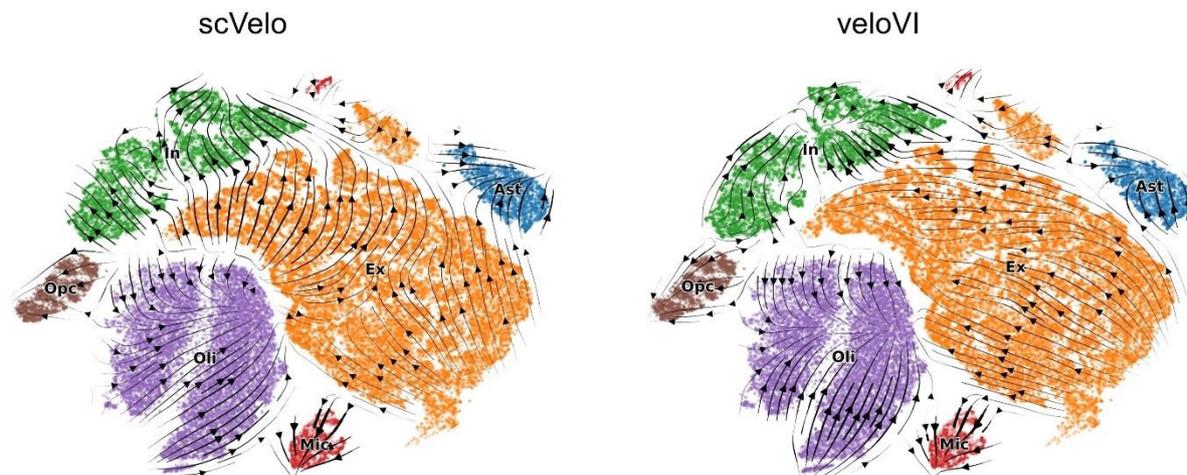


Figure 4.S2: Comparison of single-nucleus RNA velocity trajectory predictions between scVelo and veloVI. Velocities were projected on the cells across 48 subjects. Inferred directionality was concordant across cell types, except for minor differences in excitatory neurons.

## Effect of age on differential RNA velocity between controls and AD-pathology subject

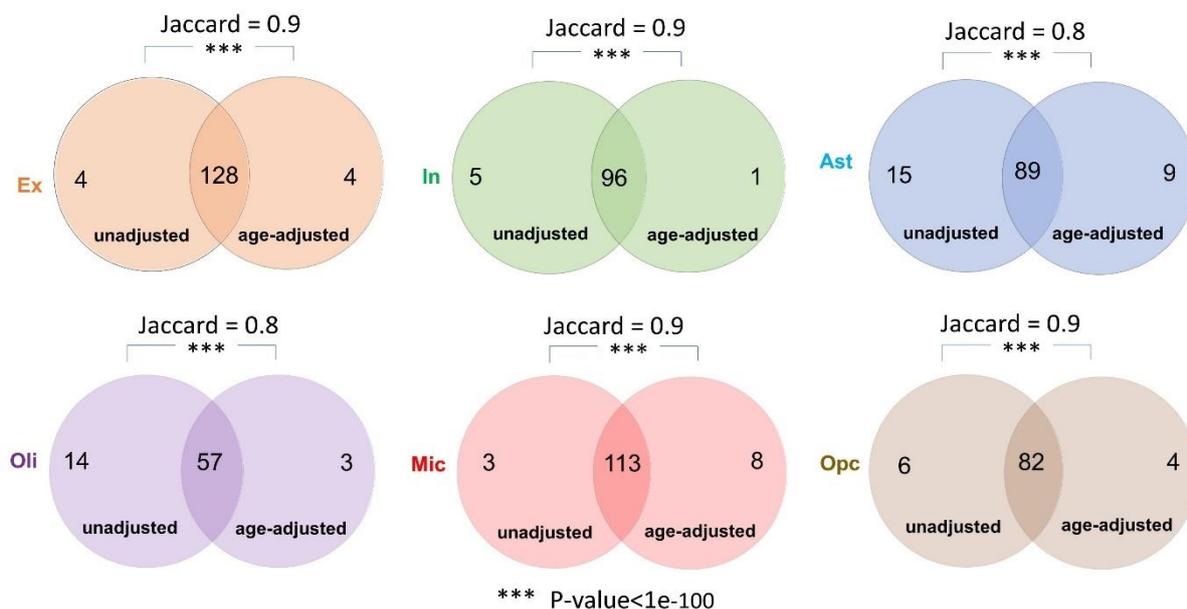


Figure 4.S3: Comparison of group differences in RNA velocity between controls and AD-pathology subjects before and after correcting for age. Jaccard similarity between the unadjusted and age-corrected analyses demonstrates the absence of confounding effect, possibly because the subjects were matched for age.



Figure 4.S5: Protein-protein interaction network of the differential velocity genes implicated in synaptic and ion channel activities in excitatory neurons. PPI network was retrieved from STRING database.

### Changes in transcriptional kinetics with AD progression

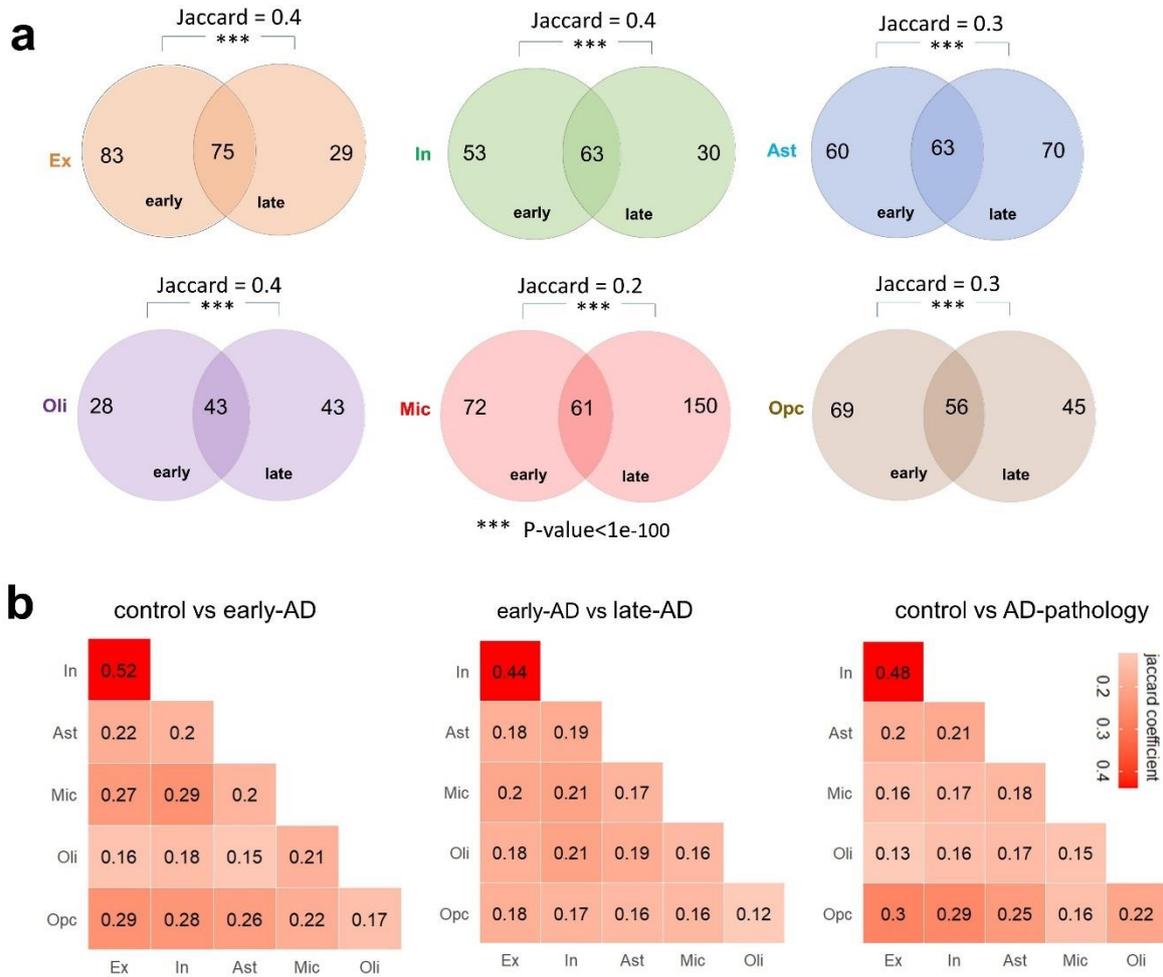


Figure 4.S6: Changes in RNA velocity with disease progression. a) Comparison of the number of genes with differential RNA velocity between “control versus early-AD analysis (left circle)” and “control versus late-AD analysis (right circle)”. b) Cell-specificity of changes in transcriptional dynamics with disease progression. First, isolated nuclei were compared between control and AD-pathology subjects to determine genes with differential RNA velocity (right). Then, the level of overlap between differential velocity genes across cell types is quantified via Jaccard similarity. The analyses were repeated for “early- versus late-AD individuals” (middle); and “control versus early-AD subjects” (left).

## Comparison of enriched pathways in male and female subjects

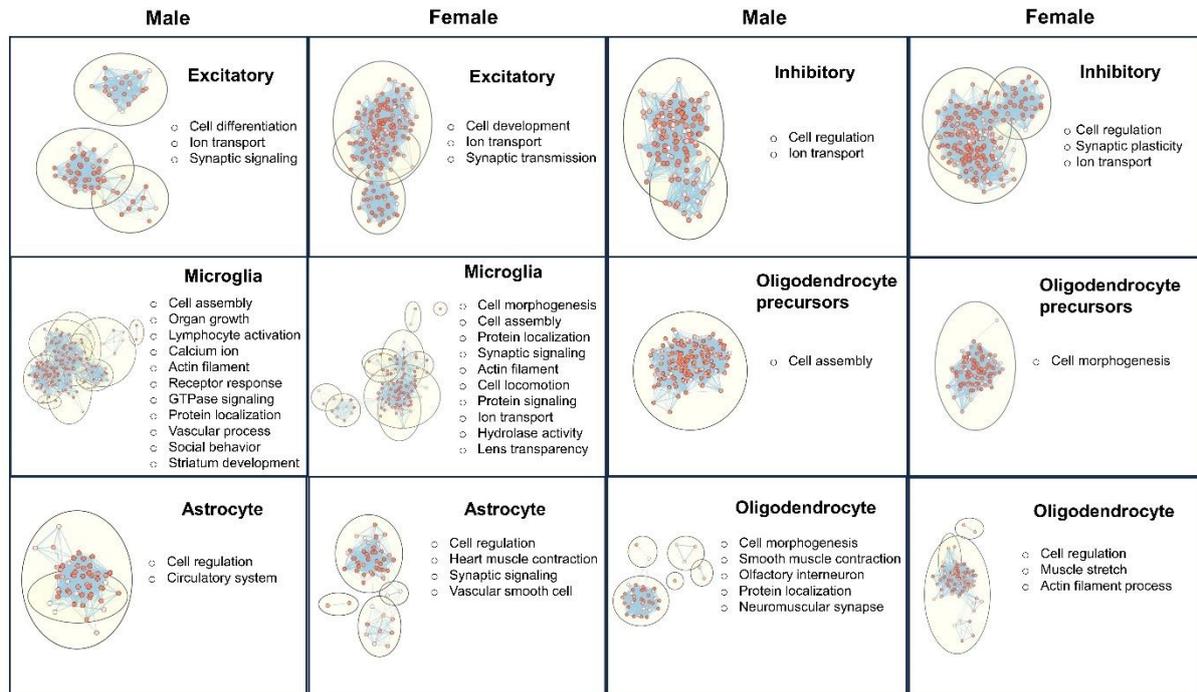


Figure 4.S7: Enriched disease pathway networks in male and female across six cell types. GO biological processes derived from differential velocity genes between controls and AD-pathology subjects are clustered into biological themes (adjusted p-value < 0.01). The biological themes are listed in increasing order of cluster size.

## Inter-cell overlap of dynamic genes correlated with AD pathology

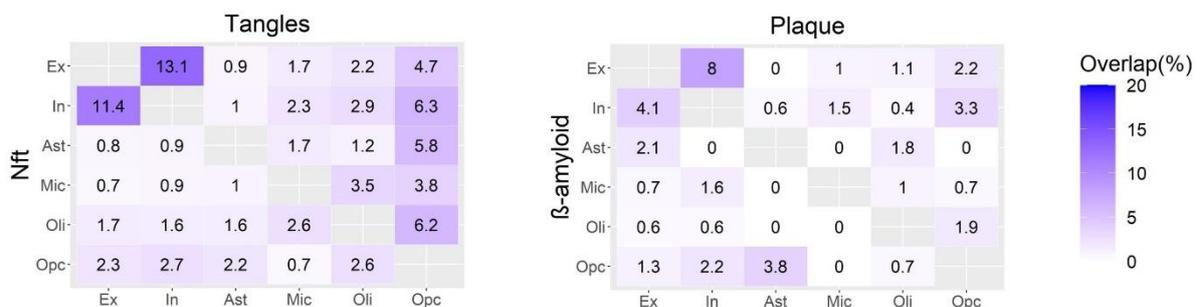


Figure 4.S8: Percentage overlap of dynamic genes associated with AD pathology in each cell type. The percentage is obtained by expressing the number of genes common to any two cell types as a fraction of the total number of unique genes in the two cell types. Two neuropathological variables are visualized simultaneously on the same plot. Only excitatory and inhibitory neurons presented a relatively high overlapping (up to 13%) in significant genes associated with the different AD neuropathological traits. The other cell types substantially differed in gene-specific changes across the phenotypes.

## Differential velocity is not confounded by post-mortem interval

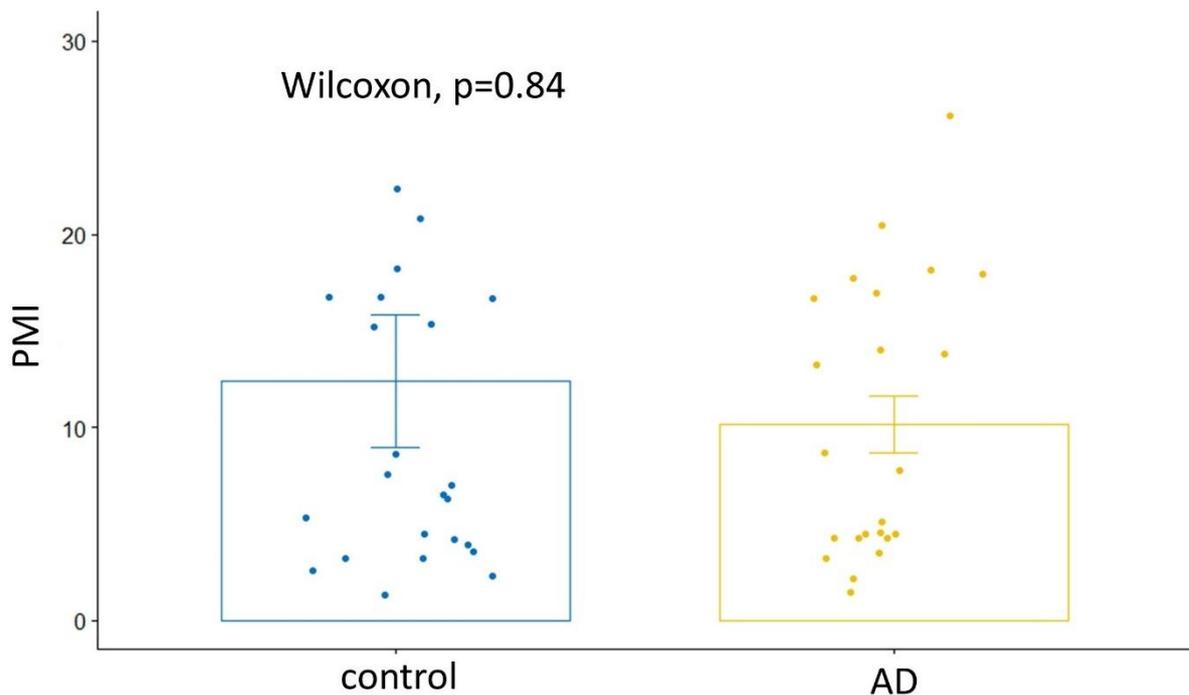


Figure 4.S9: Test for difference in post-mortem interval between the control and AD groups. Wilcoxon rank-sum test shows that there is no significant difference in the post-mortem sampling intervals, indicating that the differences captured between the two groups is related to pathological process.

## 4.8 Additional information

### 4.8.1 Data availability

Dataset-1 snRNA-seq and metadata are available at the AMP-AD portal (<https://www.synapse.org>, IDs syn18485175 and syn3157322). Dataset-2 snRNA-seq and metadata can also be downloaded from the AMP-AD portal (<https://www.synapse.org>, ID syn16780177). The raw scRNA-seq of Dataset-3 are available on the GEO (<https://www.ncbi.nlm.nih.gov/geo/>, accession code GSE135618). The data on GEO is freely accessible without registration while datasets on Synapse are available under controlled use conditions to ensure anonymity of the study participants. Hence, data use agreement and registration are required to access Dataset-1 and Dataset-2. ROSMAP data can be requested at <https://www.radc.rush.edu>.

## 4.8.2 Code availability

scVelo (v.0.2.3) and veloVI (v.0.3.0) are downloadable as python packages (see <https://scvelo.readthedocs.io/installation/> and <https://velovi.readthedocs.io/en/latest/>). The codes used for the analyses will be available with article publication at <https://neuropm-lab.com/other-pipelines>.

## 4.8.3 Acknowledgements

This project was undertaken thanks in part to the following funding awards to YIM: the Canada Research Chair tier-2, the CIHR Project Grant 2020, the Weston Family Foundation's Transformational Research in AD 2020, the Ludmer Centre for Neuroinformatics and Mental Health 2020 award, and the New Investigator start-up grant from McGill University's Healthy Brains for Healthy Lives Initiative (Canada First Research Excellence Fund). QA is partly supported by Parkinson Canada and Fonds de recherche du Québec – Santé (FRQS) Graduate Partnership Awards, and AFK is partly funded by Parkinson Canada Graduate Student Award. In addition, we used the computational infrastructure supported in part by the Brain Canada Foundation and Health Canada support to the McConnell Brain Imaging Center at the Montreal Neurological Institute. Dataset-1 and Dataset-2 (ROSMAP) were provided by the Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago. The data collection was supported through funding by NIA grants P30AG10161, P30AG72975, R01AG15819, R01AG17917, R01AG30146, R01AG36836, U01AG32984, U01AG46152, U01AG61356, the Illinois Department of Public Health, and the Translational Genomics Research Institute. Dataset-3 collection was partly supported by National Institutes of Health (NIH) awards F30 AG066418, K08 AG052648, R56 AG057528, K24 AG053435, U54 NS100717, AG023501, AG019724, the Tau Consortium, and the Bluefield Project to Cure FTD.

## References

- Adewale, Q., Khan, A. F., Carbonell, F., & Iturria-Medina, Y. (2021). Integrated transcriptomic and neuroimaging brain model decodes biological mechanisms in aging and Alzheimer's disease. *ELife*, *10*. <https://doi.org/10.7554/eLife.62589>
- Arendt, T. (2003). Synaptic plasticity and cell cycle activation in neurons are alternative effector pathways: the 'Dr. Jekyll and Mr. Hyde concept' of Alzheimer's disease or the yin and yang of neuroplasticity. *Progress in Neurobiology*, *71*(2–3), 83–248. <https://doi.org/10.1016/J.PNEUROBIO.2003.09.007>
- Belonwu, S. A., Li, Y., Bunis, D., Rao, A. A., Solsberg, C. W., Tang, A., Fragiadakis, G. K., Dubal, D. B., Oskotsky, T., & Sirota, M. (2022). Sex-Stratified Single-Cell RNA-Seq Analysis Identifies Sex-Specific and Cell Type-Specific Transcriptional Responses in Alzheimer's Disease Across Two Brain Regions. *Molecular Neurobiology*, *59*(1), 276–293. <https://doi.org/10.1007/s12035-021-02591-8>
- Bennett, D. A., Buchman, A. S., Boyle, P. A., Barnes, L. L., Wilson, R. S., & Schneider, J. A. (2018). Religious Orders Study and Rush Memory and Aging Project. *Journal of Alzheimer's Disease : JAD*, *64*(s1), S161–S189. <https://doi.org/10.3233/JAD-179939>
- Bergen, V., Lange, M., Peidli, S., Wolf, F. A., & Theis, F. J. (2020). Generalizing RNA velocity to transient cell states through dynamical modeling. *Nature Biotechnology*, *38*(12), 1408–1414. <https://doi.org/10.1038/s41587-020-0591-3>
- Blockus, H., Rolotti, S. V., Szoboszlai, M., Peze-Heidsieck, E., Ming, T., Schroeder, A., Apostolo, N., Vennekens, K. M., Katsamba, P. S., Bahna, F., Mannepalli, S., Ahlsen, G., Honig, B., Shapiro, L., de Wit, J., Losonczy, A., & Polleux, F. (2021). Synaptogenic activity of the axon guidance molecule Robo2 underlies hippocampal circuit function. *Cell Reports*, *37*(3), 109828. <https://doi.org/10.1016/j.celrep.2021.109828>
- Bray, N. L., Pimentel, H., Melsted, P., & Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology* 2016 34:5, *34*(5), 525–527. <https://doi.org/10.1038/nbt.3519>
- Cain, A., Taga, M., McCabe, C., Hekselman, I., White, C. C., Green, G., Rozenblatt-Rosen, O., Zhang, F., Yeger-Lotem, E., Bennett, D. A., Yang, H.-S., Regev, A., Menon, V., Habib, N., & Jager, P. L. De. (2020). Multi-cellular communities are perturbed in the aging human brain and with Alzheimer's disease. *BioRxiv*.
- Chen, E. Y., Tan, C. M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G. V., Clark, N. R., & Ma'ayan, A. (2013). Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics*, *14*, 128. <https://doi.org/10.1186/1471-2105-14-128>
- Couturier, C. P., Ayyadhury, S., Le, P. U., Nadaf, J., Monlong, J., Riva, G., Allache, R., Baig, S., Yan, X., Bourgey, M., Lee, C., Wang, Y. C. D., Wee Yong, V., Guiot, M.-C., Najafabadi, H., Misic, B., Antel, J., Bourque, G., Ragoussis, J., & Petrecca, K. (2020). Single-cell RNA-seq reveals that glioblastoma recapitulates a normal neurodevelopmental hierarchy. *Nature Communications* 2020 11:1, *11*(1), 1–19. <https://doi.org/10.1038/s41467-020-17186-5>

- De Strooper, B., & Karran, E. (2016). The Cellular Phase of Alzheimer's Disease. In *Cell* (Vol. 164, Issue 4, pp. 603–615). Cell Press. <https://doi.org/10.1016/j.cell.2015.12.056>
- Gayoso, A., Weiler, P., Lotfollahi, M., Klein, D., Hong, J., Streets, A., Theis, F. J., & Yosef, N. (2023). Deep generative modeling of transcriptional dynamics for RNA velocity analysis in single cells. *Nature Methods*. <https://doi.org/10.1038/s41592-023-01994-w>
- Genovese, C. R., Lazar, N. A., & Nichols, T. (2002). Thresholding of Statistical Maps in Functional Neuroimaging Using the False Discovery Rate. *NeuroImage*, 15(4), 870–878. <https://doi.org/10.1006/NIMG.2001.1037>
- Gerrits, E., Heng, Y., Boddeke, E. W. G. M., & Eggen, B. J. L. (2020). Transcriptional profiling of microglia; current state of the art and future perspectives. *Glia*, 68(4), 740–755. <https://doi.org/10.1002/glia.23767>
- Gorin, G., & Pachter, L. (2021). Length Biases in Single-Cell RNA Sequencing of pre-mRNA. *BioRxiv*, 2021.07.30.454514. <https://doi.org/10.1101/2021.07.30.454514>
- Habib, N., Avraham-Davidi, I., Basu, A., Burks, T., Shekhar, K., Hofree, M., Choudhury, S. R., Aguet, F., Gelfand, E., Ardlie, K., Weitz, D. A., Rozenblatt-Rosen, O., Zhang, F., & Regev, A. (2017). Massively parallel single-nucleus RNA-seq with DroNc-seq. *Nature Methods*, 14(10), 955–958. <https://doi.org/10.1038/NMETH.4407>
- Joseph, C., Mangani, A. S., Gupta, V., Chitranshi, N., Shen, T., Dheer, Y., Kb, D., Mirzaei, M., You, Y., Graham, S. L., & Gupta, V. (2020). Cell Cycle Deficits in Neurodegenerative Disorders: Uncovering Molecular Mechanisms to Drive Innovative Therapeutic Development. *Aging and Disease*, 11(4), 946–966. <https://doi.org/10.14336/AD.2019.0923>
- Kaneko, N., Marín, O., Koike, M., Hirota, Y., Uchiyama, Y., Wu, J. Y., Lu, Q., Tessier-Lavigne, M., Alvarez-Buylla, A., Okano, H., Rubenstein, J. L. R., & Sawamoto, K. (2010). New neurons clear the path of astrocytic processes for their rapid migration in the adult brain. *Neuron*, 67(2), 213–223. <https://doi.org/10.1016/j.neuron.2010.06.018>
- Kanton, S., Boyle, M. J., He, Z., Santel, M., Weigert, A., Sanchís-Calleja, F., Guijarro, P., Sidow, L., Fleck, J. S., Han, D., Qian, Z., Heide, M., Huttner, W. B., Khaitovich, P., Pääbo, S., Treutlein, B., & Camp, J. G. (2019). Organoid single-cell genomic atlas uncovers human-specific features of brain development. *Nature*, 574(7778), 418–422. <https://doi.org/10.1038/s41586-019-1654-9>
- Khan, A. F., Adewale, Q., Baumeister, T. R., Carbonell, F., Zilles, K., Palomero-Gallagher, N., Iturria-Medina, Y., & Initiative, the A. D. N. (2021). Personalized brain models identify neurotransmitter receptor changes in Alzheimer's disease. *Brain*. <https://doi.org/10.1093/brain/awab375>
- Kuleshov, M. V., Jones, M. R., Rouillard, A. D., Fernandez, N. F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S. L., Jagodnik, K. M., Lachmann, A., McDermott, M. G., Monteiro, C. D., Gundersen, G. W., & Ma'ayan, A. (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Research*, 44(W1), W90–7. <https://doi.org/10.1093/nar/gkw377>
- La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber,

- K., Kastri, M. E., Lönnerberg, P., Furlan, A., Fan, J., Borm, L. E., Liu, Z., van Bruggen, D., Guo, J., He, X., Barker, R., Sundström, E., Castelo-Branco, G., ... Kharchenko, P. V. (2018). RNA velocity of single cells. *Nature*, *560*(7719), 494–498. <https://doi.org/10.1038/s41586-018-0414-6>
- Lake, B. B., Codeluppi, S., Yung, Y. C., Gao, D., Chun, J., Kharchenko, P. V., Linnarsson, S., & Zhang, K. (2017). A comparative strategy for single-nucleus and single-cell transcriptomes confirms accuracy in predicted cell-type expression from nuclear RNA. *Scientific Reports*, *7*(1), 6031. <https://doi.org/10.1038/s41598-017-04426-w>
- Lau, S. F., Cao, H., Fu, A. K. Y., & Ip, N. Y. (2020). Single-nucleus transcriptome analysis reveals dysregulation of angiogenic endothelial cells and neuroprotective glia in Alzheimer's disease. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(41), 25800–25809. <https://doi.org/10.1073/pnas.2008762117>
- Leng, K., Li, E., Eser, R., Piergies, A., Sit, R., Tan, M., Neff, N., Li, S. H., Rodriguez, R. D., Suemoto, C. K., Leite, R. E. P., Ehrenberg, A. J., Pasqualucci, C. A., Seeley, W. W., Spina, S., Heinsen, H., Grinberg, L. T., & Kampmann, M. (2021). Molecular characterization of selectively vulnerable neurons in Alzheimer's disease. *Nature Neuroscience*, *24*(2), 276–287. <https://doi.org/10.1038/s41593-020-00764-7>
- Lo Giudice, Q., Leleu, M., La Manno, G., & Fabre, P. J. (2019). Single-cell transcriptional logic of cell-fate specification and axon guidance in early-born retinal neurons. *Development (Cambridge, England)*, *146*(17). <https://doi.org/10.1242/dev.178103>
- Marsh, B., & Blesch, R. (2020). Single nuclei RNA-seq of mouse placental labyrinth development. *ELife*, *9*, 1–27. <https://doi.org/10.7554/ELIFE.60266>
- Maruszak, A., & Zekanowski, C. (2011). Mitochondrial dysfunction and Alzheimer's disease. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *35*(2), 320–330. <https://doi.org/10.1016/J.PNPBP.2010.07.004>
- Mathys, H., Davila-Velderrain, J., Peng, Z., Gao, F., Mohammadi, S., Young, J. Z., Menon, M., He, L., Abdurrob, F., Jiang, X., Martorell, A. J., Ransohoff, R. M., Hafler, B. P., Bennett, D. A., Kellis, M., & Tsai, L. H. (2019). Single-cell transcriptomic analysis of Alzheimer's disease. *Nature*, *570*(7761), 332–337. <https://doi.org/10.1038/s41586-019-1195-2>
- Melsted, P., Boeshaghi, A. S., Liu, L., Gao, F., Lu, L., Min, K. H. (Joseph), da Veiga Beltrame, E., Hjärleifsson, K. E., Gehring, J., & Pachter, L. (2021). Modular, efficient and constant-memory single-cell RNA-seq preprocessing. *Nature Biotechnology* *2021* *39*:7, *39*(7), 813–818. <https://doi.org/10.1038/s41587-021-00870-2>
- Meyer, K., Feldman, H. M., Lu, T., Drake, D., Lim, E. T., Ling, K.-H., Bishop, N. A., Pan, Y., Seo, J., Lin, Y.-T., Su, S. C., Church, G. M., Tsai, L.-H., & Yankner, B. A. (2019). REST and Neural Gene Network Dysregulation in iPSC Models of Alzheimer's Disease. *Cell Reports*, *26*(5), 1112–1127.e9. <https://doi.org/10.1016/j.celrep.2019.01.023>
- Nikolakopoulou, A. M., Koeppen, J., Garcia, M., Leish, J., Obenaus, A., & Ethell, I. M. (2016). Astrocytic Ephrin-B1 Regulates Synapse Remodeling Following Traumatic Brain Injury. *ASN Neuro*, *8*(1), 1–18. <https://doi.org/10.1177/1759091416630220>

- Olah, M., Menon, V., Habib, N., Taga, M. F., Ma, Y., Yung, C. J., Cimpean, M., Khairallah, A., Coronas-Samano, G., Sankowski, R., Grün, D., Kroshilina, A. A., Dionne, D., Sarkis, R. A., Cosgrove, G. R., Helgager, J., Golden, J. A., Pennell, P. B., Prinz, M., ... De Jager, P. L. (2020). Single cell RNA sequencing of human microglia uncovers a subset associated with Alzheimer's disease. *Nature Communications*, *11*(1), 1–18. <https://doi.org/10.1038/s41467-020-19737-2>
- Pan, X., Zhang, H., Xu, D., Chen, J., Chen, W., Gan, S., Qu, F., Chu, C., Cao, J., Fan, Y., Song, X., Ye, J., Zhou, W., & Cui, X. (2020). Identification of a novel cancer stem cell subpopulation that promotes progression of human fatal renal cell carcinoma by single-cell RNA-seq analysis. *International Journal of Biological Sciences*, *16*(16), 3149. <https://doi.org/10.7150/IJBS.46645>
- Park, H., Choi, Y., Jung, H., Kim, S., Lee, S., Han, H., Kweon, H., Kang, S., Sim, W. S., Koopmans, F., Yang, E., Kim, H., Smit, A. B., Bae, Y. C., & Kim, E. (2020). Splice-dependent trans-synaptic PTP $\delta$ -IL1RAPL1 interaction regulates synapse formation and non-REM sleep. *The EMBO Journal*, *39*(11), e104150. <https://doi.org/10.15252/emj.2019104150>
- Ramos-Brossier, M., Montani, C., Lebrun, N., Gritti, L., Martin, C., Seminatore-Nole, C., Toussaint, A., Moreno, S., Poirier, K., Dorseuil, O., Chelly, J., Hackett, A., Gecz, J., Bieth, E., Faudet, A., Heron, D., Frank Kooy, R., Loeys, B., Humeau, Y., ... Billuart, P. (2015). Novel IL1RAPL1 mutations associated with intellectual disability impair synaptogenesis. *Human Molecular Genetics*, *24*(4), 1106–1118. <https://doi.org/10.1093/hmg/ddu523>
- Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., & Vilo, J. (2019). g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Research*, *47*(W1), W191–W198. <https://doi.org/10.1093/NAR/GKZ369>
- Santiago, C. P., Gimmen, M. Y., Lu, Y., McNally, M. M., Duncan, L. H., Creamer, T. J., Orzolek, L. D., Blackshaw, S., & Singh, M. S. (2023). Comparative Analysis of Single-cell and Single-nucleus RNA-sequencing in a Rabbit Model of Retinal Detachment-related Proliferative Vitreoretinopathy. *Ophthalmology Science*, *3*(4), 100335. <https://doi.org/10.1016/J.XOPS.2023.100335>
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., & Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research*, *13*(11), 2498–2504. <https://doi.org/10.1101/GR.1239303>
- Szklarczyk, D., Gable, A. L., Nastou, K. C., Lyon, D., Kirsch, R., Pyysalo, S., Doncheva, N. T., Legeay, M., Fang, T., Bork, P., Jensen, L. J., & von Mering, C. (2021). The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Research*, *49*(D1), D605–D612. <https://doi.org/10.1093/nar/gkaa1074>
- Wolfien, M., Galow, A.-M., Müller, P., Bartsch, M., Brunner, R. M., Goldammer, T., Wolkenhauer, O., Hoeflich, A., & David, R. (2020). Single-Nucleus Sequencing of an Entire Mammalian Heart: Cell Type Composition and Velocity. *Cells*, *9*(2). <https://doi.org/10.3390/cells9020318>

Yang, C.-P., Gilley, J. A., Zhang, G., & Kernie, S. G. (2011). ApoE is required for maintenance of the dentate gyrus neural progenitor pool. *Development (Cambridge, England)*, *138*(20), 4351–4362. <https://doi.org/10.1242/dev.065540>

Zuguang, G., Roland, E., & Matthias, S. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics (Oxford, England)*, *32*(18), 2847–2849. <https://doi.org/10.1093/BIOINFORMATICS/BTW313>

## **Chapter 5: Discussion**

### **5.1 Summary of findings**

Using mechanistic data-driven approaches, this thesis aimed to advance the understanding of the multifactorial and cellular alterations of the brain in AD and PD, and how these changes drive patient symptoms and clinical heterogeneity. Specifically, we developed a novel mathematical model called Gene Expression Multifactorial Causal Model (GE-MCM), which was used to incorporate multiple biomarkers to study these two progressive disorders. Chapter 1 applied the GE-MCM to uncover the genes and biological mechanisms underlying normal aging and AD. Chapter 2 demonstrated the disease-agnostic utility of the model by applying it to PD to unravel the biological substrates of clinical symptoms and physical activity, and identify putative disease-modifying therapeutic targets. Chapter 3 applied another mechanistic model, called RNA velocity, to study neurodegenerative diseases for the first time, uncovering the dynamical molecular changes underlying AD progression across different brain cell types. The implications of these studies are discussed below.

#### **5.1.1 GE-MCM for aging and AD**

The main goal of this study was to develop a novel mechanistic mathematical model of aging and neurodegenerative diseases. In an unprecedented way, the GE-MCM allows the incorporation of multiple biomarkers at different spatial resolutions. The multiscale interactions between the biomarkers can be studied, to provide mechanistic insights into the complexity of aging and disease progression. Moreover, the model is applied at the individual level to capture disentangle disease heterogeneity and develop personalized treatment.

By combining whole-brain gene expression with six neuroimaging modalities, we identified genes driving brain reorganization and clinical symptoms in AD. We also uncovered few genes underlying healthy aging. In addition, the possible disease-related factor modulated

by these genes are revealed, providing a plausible explanation for the pathological interactions that drive the disease process. Interestingly, some of these gene-neuroimaging interactions have been reported by earlier experimental studies. Our study therefore provides novel associations that can be functionally validated.

Another goal of this project was to conceptualise the relationship between normal aging and AD. We found that the number of genes associated with AD were far more than those of normal aging. Interestingly, only one gene (*LSM6*) was common to the two processes. Nevertheless, as many genes might be implicated in a single biological pathway, we further investigated if the pathways underlying normal aging and AD are distinct. Our findings demonstrated both aging and AD share some biological pathways including oxidative stress, immune/inflammatory response, and G-protein couple receptor signalling. However, more pathways are implicated in AD, demonstrating that the disease is a more complex process than aging. Whether AD is a distinct entity or a continuum of aging is a subject of an ongoing debate (Franceschi et al., 2018). However, if AD is a distinct process, alterations in cognitive and neuropsychiatric functions at prodromal or early disease stage might be mistakenly attributed to normal aging. In any case, our findings have implications for identifying patients at risk of moving from MCI to AD. Indeed, therapeutic treatments and lifestyle interventions that target the common biological pathways of normal aging and AD can help reduce the risk of disease progression or time of clinical onset.

### **5.1.2 GE-MCM for PD**

By also using GE-MCM to study PD, we showed that this novel method is generally applicable to progressive diseases. Specifically, we studied the biological mechanisms underlying different motor and non-motor symptoms in PD, identified the molecular substrates of physical activity, and demonstrated the translational utility of the model for drug discovery.

One of the original findings of this project was that PD overlaps with cancer and infection pathways, highlighting the systemic interaction of PD with other disorders. Our analysis also transcends traditional discrete methods of disease subtyping, wherein defining accurate cut-off points and homogeneity of subtypes remains a challenge. Instead, we employed continuous approach to disentangle disease heterogeneity by identifying the biological mechanisms that underlie distinct combinations of clinical symptom profiles (including motor, non-motor, psychiatric and autonomic dysfunctions). To further translate symptom profiles to treatment needs, we uncovered the protein-protein interaction (PPI) networks underlying the biological mechanisms. Interestingly, 5 out of the 6 hub genes underlying these PPI networks have been previously associated with PD from different study cohorts. Thus, in addition to identifying a novel PD hub gene, our study links the hub genes to different PD clinical symptom profiles. As hub genes interact with many other genes and play critical roles in biological processes, the hub genes identified in our study can guide biomarker and treatment identification by matching the corresponding symptom profiles with real-life clinical cases.

Another important part of this project was to investigate the biological mechanisms associating PD with physical activity. We discovered that physical activity is associated with PD through two principal pathways, namely, cholesterol biosynthesis and immune response. Specifically, the cholesterol biosynthesis pathway may act as biological substrate for leisure activities while immune response was more related to household activities. Work-related activity seems to play moderate roles in both pathways. Even though a bidirectional relationship exists between physical activity and PD, the results hint at possible biological pathways that could be modulated to enhance the benefits of physical activity in PD.

Lastly, we demonstrated the potential use of GE-MCM for identifying disease-modifying treatments. Through personalized *in silico* gene perturbation, we uncovered putative

drug targets for PD. Interestingly, many of these identified drugs are associated with dopamine agonists, including Levodopa which is the mainstay of PD treatment. Other drugs are associated with infection, inflammation, and insomnia. While some of these drugs are symptomatic, the results point to a promising direction of drug repurposing and refinement for slowing down or halting disease progression.

### **5.1.3 Single-cell RNA velocity changes in AD**

The aim of this project was to investigate dynamic molecular dysregulations underlying AD. By transcending the traditional static gene expression analyses, we applied a mechanistic model called RNA velocity to single-nucleus RNA data from different brain cell types.

Compared to single cells, it is more practical to isolate single nuclei from postmortem brain tissues. However, RNA velocity was originally developed for single-cell data and its validity for single-nucleus data is debated. We demonstrated that RNA velocity can also be calculated for single-nucleus RNA-seq and the estimates are concordant with those obtained from single-cell RNA-seq. We then carried out differential analysis between controls and AD subjects by separately using static gene expression and dynamic RNA velocity. We found that the genes underlying differential velocity are largely different from those underlying static gene expression. Importantly, velocity-associated genes are mainly involved in cell organization and synaptic processes. The latter finding was replicated in an independent dataset, demonstrating that the results are generalizable and stable.

Investigating the sex-dimorphic differences in RNA velocity changes revealed that more genes are dysregulated in females compared to males across all cell types. Nevertheless, the dysregulated pathways between both sexes are largely similar except for lymphocyte activation and vascular process which are pronounced in the microglia of males. We also examined the dependence of RNA velocity changes with disease stage and observed increased

number of affected genes in microglia, astrocytes, and oligodendrocytes. Thus, AD is characterized by progressive immune dysregulation with disease advancement.

Further analysis revealed that the RNA velocity changes are highly associated with neuropathological burden, and these associations are cell-type specific. Hence, moving beyond neurocentric view and including glial cell types is crucial for a better understanding of AD pathogenesis. Finally, we observed accelerated cell changes in AD subjects compared to controls, suggesting that early depletion of cells could be an underlying driver of neurodegeneration.

## **5.2 Limitations**

### **5.2.1 GE-MCM for aging, AD and PD**

Despite the ability of the novel GE-MCM model to accurately recover the aging and neurodegenerative processes being modelled, some limitations still need to be acknowledged. First, the ADNI and PPMI cohorts do not have patient-specific gene expression data. Hence, we derived a brain-wide gene expression data by combining mRNA values from six healthy adult brains. Notably, the gene expression data is obtained for a single time point. However, the model was fitted for individual patient using temporally varying longitudinal neuroimage modalities. The parameters obtained after model fitting are assumed to reflect the gene-specific deformations required to fit the data, hence quantitative measures of the individual dysregulation.

Second, one of the main challenges of longitudinal studies is the presence of missing data. Correspondingly, the GE-MCM model required at least 4 neuroimaging modalities measured across at least three time points. As a result, subjects who did not meet this criterion were excluded. For the remaining subjects, uniformity in the number of time points across all six neuroimaging modalities was ensured by imputation. Even though the imputation was validated, using actual non-imputed data could improve the model fitting results.

Third, even though the GE-MCM incorporated hundreds of genes that have been previously shown to capture many biological functions, the human genome has over 20,000 genes. It is therefore possible that some genes with important roles in aging and disease processes are missed. However, the dimensions of the GE-MCM model increases by six folds for each additional gene included. To incorporate more genes while maintaining a stable model fitting, a greater number of subjects will be required.

Fourth, most of our analyses assume linear dependencies. For example, the GE-MCM linearly models the interaction between genes and neuroimaging-derived measures. Although the actual biological processes (i.e., gene and neuroimaging) could be interacting nonlinearly, it is difficult to systematically find an analytic solution to a nonlinear model. Furthermore, by using a linear multivariate mapping, we examined how the interactions between genes and neuroimaging measures affect clinical changes in aging and neurodegeneration. While it is possible to include non-linearity as commonly done in regression analysis, studying nonlinear interdependencies would require separate addition of linear and non-linear terms. This is almost impractical given the high dimensionality of the model. We also assumed that the changes in clinical measures over time follows a linear trend. Studies have however shown that the trajectories of clinical symptoms could be non-linear (Bhagwat et al., 2018; Jack et al., 2012). Fitting non-linear models such as sigmoid or polynomial functions to the clinical scores is worth considering.

Fifth, our analysis of normal aging used ADNI healthy participants who remained clinically healthy for an average of 7.8 years. It was impossible to immediately ascertain if these subjects will eventually develop MCI or AD. Nevertheless, the common duration of follow-up of clinical trials is 4 years from healthy state to MCI conversion, and up to 3 years from MCI to AD (Pang et al., 2023). We therefore believe that remaining clinically healthy for about 8 years is sufficient to model normal aging for practical purpose.

Sixth, while studying the relationship between PD and physical activity, we didn't not disentangle the direction of causality. Indeed, physical activity can ameliorate PD symptoms. Conversely, aggravating PD symptoms can impede patients from carrying out physical activity. Thus, we interpreted our findings in the light of mediation, identifying immune response and cholesterol metabolism as important mechanism mediating the link between PD and physical activity.

Lastly, while the GE-MCM was interpreted as capturing the direct influence of genes on different biological processes, it should be noted that genes function after being translated to proteins. Moreover, the genes/proteins rarely function in isolation but interact with one another to influence biological processes. Even though the GE-MCM incorporated the effect of individual genes separately, the landmark genes used in these studies represent the centers of various interacting proteins.

### **5.2.2 RNA velocity changes in AD**

The main limitation of Chapter 4 of this study lies with the RNA velocity estimation itself. RNA velocity is nascent and its methods for estimation are continuously being improved upon (Bergen et al. 2021; Zheng et al. 2023). Importantly, RNA velocity was originally developed for single-nucleus data where the nucleus and cytoplasm are intact. However, post-mortem tissues are frozen and extracting whole cell is challenging. Moreover, dissociating whole cell from the brain is difficult because the intense enzymatic activity required can interfere with cell recovery. Hence, single nuclei provide an amenable alternative for sequencing post-mortem brain tissues. The accuracy of single nuclei for RNA velocity calculation is debated due to the loss of cytoplasmic content. However, our validation analysis showed a high concordance between RNA velocity values estimated from single cells and nuclei of microglia of some patients who underwent surgery. A recent animal study also showed concordant results between the RNA velocity trajectories of single nuclei and single

cells in retina (Santiago et al., 2023). Other studies obtained biologically meaningful results by using single-nucleus RNA velocity to infer the developmental process of mouse embryo and human heart (Marsh & Blelloch, 2020; Wolfien et al., 2020). More insights into the validity of RNA velocity for single nuclei will emerge as more complex modelling methods are developed, such as replacing mRNA degradation of single cells with nuclear export of single nucleus.

The second limitation lies in the difference between the timescales of RNA and neurodegeneration. The life cycle of RNA is a few days, which is much shorter than the timescale of neurodegeneration. Our study provided the first use case of RNA velocity for this slowly evolving degenerative process. We believe that using subjects with varied levels of neuropathology provided an opportunity to globally capture the association between the timescales of transcriptional dynamics and neurodegeneration. Barring experimental limitations, a single subject could be sampled at several post-mortem intervals to validate the stability of RNA velocity within that subject. Nevertheless, we confirmed that post-mortem sampling intervals is not associated with RNA velocity differences between healthy and AD subjects, despite that death process could cause transcriptional changes.

## **5.3 Future direction**

### **5.3.1 Integration of peripheral biomarkers**

The GE-MCM offers a flexible formulation that can allow integration of other biomarkers. In lieu of generating whole-brain gene expression which can be expensive without interpolation, blood-based gene expression can be used. Interestingly, animal and human studies have shown promising concordance between peripheral and brain gene expression patterns, suggesting that blood-based gene expression can serve as surrogate for gene expression in the brain (Iturria-Medina et al., 2020; Jasinska et al., 2009; Sullivan et al., 2006; Witt et al., 2013). Moreover, developing non-invasive biomarkers can offer a fast screening

tool for large scale research and routine clinical practice (Delaby et al., 2023; Leuzy et al., 2022). Validating a peripheral biomarker in the context of GE-MCM is a viable approach.

### **5.3.2 Incorporating cell-type specific changes**

The advent of single-cell sequencing technologies has provided an unprecedented opportunity to study diverse brain cell types and transcend the neuro-centric view of neurodegenerative diseases. However, sequencing the whole brain at single-cell level is cost prohibitive. Consequently, ingenious methods have been derived to infer and validate cell type proportions from bulk RNA sequencing (McKenzie et al., 2018; Newman et al., 2019). Importantly, a recent study derived whole-brain gene expression patterns from Allen Human Brain Atlas by interpolation, and subsequently applied cell type deconvolution method to infer the proportion of neural and glial cells in the healthy brain (Pak et al., 2023). These cell type proportions could replace gene expression values in the GE-MCM model to elucidate how each cell type drives disease progression. An alternative approach can incorporate cell-cell communication between cell types to understand how alterations in cell signalling drive disease (Wilk et al., 2024). Targeting the ligand-receptor complexes underlying pathogenic signalling in these diverse cell types holds promise.

### **5.3.3 Analysis of comorbidity**

Co-pathology in neurodegenerative diseases is a norm rather than an exception, especially with increased age.  $\beta$ -amyloid and tau are observed in PD patients (Mihaescu et al. 2022; Bellomo et al. 2024), and  $\alpha$ -synuclein can be found in AD subjects (Robinson et al. 2018). Using similar neuroimaging modalities obtained from healthy and AD subjects in the ADNI cohort, we identified the common and distinct biological pathways underlying normal aging and AD. Some of these neuroimaging modalities are absent in the PPMI cohort of PD patients. Nevertheless, we found that inflammation pathway is common to both aging, AD and

PD. Aligning the neuroimaging modalities will facilitate a precise comparison between AD and PD for understanding comorbidity.

#### **5.3.4 Application of GE-MCM to other progressive diseases**

The GE-MCM provides a disease-agnostic framework for modelling complex progressive diseases. Here, we have successfully applied it to aging, AD and PD. Interestingly, an macroscopic-scale multifactorial causal modelling approach of neuroimaging modalities only was used to uncover the distinct patterns of disease onset and progression in the genetic subtypes of frontotemporal dementia (McCarthy, 2022). GE-MCM can directly incorporate genetic data to provide comprehensive mechanistic insights into the pathogenesis of these various subtypes. It can also be applied to other neurodegenerative diseases such as Huntington's disease, amyotrophic lateral sclerosis, etc.

#### **5.3.5 Comparison between subtype and spectrum**

Subtyping helps disentangle heterogeneity in complex neurodegenerative diseases. Often, discrete subtyping methods produce subtypes wherein some variations are still observed within groups. Whether subtypes really represent distinct entities, or they are just variations of a continuous spectrum remains an open question. In this thesis, we used the GE-MCM to uncover the molecular underpinnings of different clinical profiles of PD patients. This quasi-continuous approach allowed us to disentangle symptomatic and molecular heterogeneity despite having a handful of subjects. One of the requirements of stable and reliable subtyping is having many subjects that capture as much disease heterogeneity as possible. As more subjects become available, future studies can compare the symptom profiles obtained via this quasi-continuous approach and traditional subtyping methods by using the same gene-imaging parameters (obtained from GE-MCM optimization on each individual data) as inputs.

### 5.3.6 Sex-specific analyses

Sex plays significant roles on disease risk, symptom onset and manifestation, and response to treatment (Philippe de Souza Ferreira et al., 2022; Pinares-Garcia et al., 2018; Zalewska et al., 2023). Our analysis of sex-specific changes in RNA velocity in AD revealed that females experience broader dynamical gene changes compared males. However, due to high data dimensionality and limited number of subjects, we could not investigate sex-related effects in the GE-MCM analyses. As model fitting of the GE-MCM is done at individual level, downstream analyses can stratify subjects by sex before studying the associations between the model output parameters and clinical evaluations or physical activity. In addition, results from individual *in silico* perturbation can be combined separately by sex to identify how different drugs might be preferentially beneficial to each sex.

### 5.3.7 Translation to clinical practice

Beyond advancing mechanistic understanding of neurogenerative and aging processes, GE-MCM can be used as a tool for drug target discovery and patient selection in clinicals. First, the interactions that we observed between genes and several neuroimaging-derived biological measures can be validated experimentally. Furthermore, the different protein-protein interactions and hub genes identified can be tested against real-life patients that present matching symptom profiles. Lastly, the model can serve as a tool for *in silico* gene perturbation to select a small group of molecules for testing. Evaluations from these clinical applications can help finetune and update the model.

## 5.4 Conclusion

In conclusion, we demonstrated the utility of a novel dynamical system model for mechanistic understanding of healthy aging and neurodegenerative diseases. We found that AD is a more complex process compared to aging, even though they shared some common

mechanisms. We also found that inflammation is a common process to aging, AD and PD. Further analysis showed that distinct molecular profiles underlie different PD clinical profiles, and this profiling can be done in a quasi-continuous manner instead of discrete subtyping. In addition to mechanistic insights, *in silico* gene perturbation at the patient level revealed putative PD drugs, demonstrating the clinical utility of the model.

Leveraging another dynamical model, we showed that studying transcriptional dynamics instead of static snapshot of gene expression could offer novel and complementary insights into neurodegeneration. We found that dysregulation in cell developmental and synaptic processes can drive pathological changes in AD across neurons and glia. In addition, accelerated cell changes resulting in depleted progenitor cell pools could engender pathological cascade in AD.

Continuous integration of multiple measures of disease-related processes at different spatiotemporal resolutions will facilitate the understanding of complex and heterogenous disease processes, the identification of inclusive biomarkers, and the development of efficient tools for clinical trials and *in silico* drug discovery.

## Bibliography

- Abi Nader, C., Ayache, N., Robert, P., & Lorenzi, M. (2020). Monotonic Gaussian Process for spatio-temporal disease progression modeling in brain imaging data. *NeuroImage*, *205*, 116266. <https://doi.org/10.1016/J.NEUROIMAGE.2019.116266>
- Agosta, F., Canu, E., Stefanova, E., Sarro, L., Tomić, A., Špica, V., Comi, G., Kostić, V. S., & Filippi, M. (2014). Mild cognitive impairment in Parkinson's disease is associated with a distributed pattern of brain white matter damage. *Human Brain Mapping*, *35*(5), 1921–1929. <https://doi.org/10.1002/HBM.22302>
- Alawode, D. O. T., Fox, N. C., Zetterberg, H., & Heslegrave, A. J. (2022). Alzheimer's Disease Biomarkers Revisited From the Amyloid Cascade Hypothesis Standpoint. *Frontiers in Neuroscience*, *16*, 837390. <https://doi.org/10.3389/FNINS.2022.837390/BIBTEX>
- Annese, V., Barcia, C., Ros-Bernal, F., Gómez, A., Ros, C. M., De Pablos, V., Fernández-Villalba, E., De Stefano, M. E., & Herrero, M. T. (2013). Evidence of oligodendrogliosis in 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP)-induced Parkinsonism. *Neuropathology and Applied Neurobiology*, *39*(2), 132–143. <https://doi.org/10.1111/J.1365-2990.2012.01271.X>
- Armstrong, R. (2020). What causes neurodegenerative disease? *Folia Neuropathologica*, *58*(2), 93–112. <https://doi.org/10.5114/fn.2020.96707>
- Ayodele, T., Rogaeva, E., Kurup, J. T., Beecham, G., & Reitz, C. (2021). Early-Onset Alzheimer's Disease: What Is Missing in Research? *Current Neurology and Neuroscience Reports*, *21*(2), 1–10. <https://doi.org/10.1007/S11910-020-01090-Y/FIGURES/1>
- Baumann, N., & Pham-Dinh, D. (2001). Biology of oligodendrocyte and myelin in the mammalian central nervous system. *Physiological Reviews*, *81*(2), 871–927. <https://doi.org/10.1152/PHYSREV.2001.81.2.871/ASSET/IMAGES/LARGE/9J0210133008.JPEG>
- Bekris, L. M., Yu, C. E., Bird, T. D., & Tsuang, D. W. (2010). Review Article: Genetics of Alzheimer Disease. <Http://Dx.Doi.Org/10.1177/0891988710383571>, *23*(4), 213–227. <https://doi.org/10.1177/0891988710383571>
- Bellomo, G., Toja, A., Paolini Paoletti, F., Ma, Y., Farris, C. M., Gaetani, L., Salvadori, N., Chiasserini, D., Wojdała, A. L., Concha-Marambio, L., & Parnetti, L. (2024). Investigating alpha-synuclein co-pathology in Alzheimer's disease by means of cerebrospinal fluid alpha-synuclein seed amplification assay. *Alzheimer's & Dementia*, *20*(4), 2444–2452. <https://doi.org/10.1002/ALZ.13658>
- Bergen, V., Soldatov, R. A., Kharchenko, P. V., & Theis, F. J. (2021). RNA velocity—current challenges and future perspectives. *Molecular Systems Biology*, *17*(8), 10282. <https://doi.org/10.15252/MSB.202110282/ASSET/0F97EEBF-57D7-4DAC-9179-FA4F87CFA9B6/ASSETS/GRAPHIC/MSB202110282-FIG-0004-M.PNG>
- Bertram, L., & Tanzi, R. E. (2019). Alzheimer disease risk genes: 29 and counting. *Nature Reviews Neurology* *2019 15:4*, *15*(4), 191–192. <https://doi.org/10.1038/s41582-019->

- Bhagwat, N., Viviano, J. D., Voineskos, A. N., & Chakravarty, M. M. (2018). Modeling and prediction of clinical symptom trajectories in Alzheimer's disease using longitudinal data. *PLOS Computational Biology*, *14*(9), e1006376. <https://doi.org/10.1371/JOURNAL.PCBI.1006376>
- Blinkouskaya, Y., Caçoilo, A., Gollamudi, T., Jalalian, S., & Weickenmeier, J. (2021). Brain aging mechanisms with mechanical manifestations. *Mechanisms of Ageing and Development*, *200*, 111575. <https://doi.org/10.1016/J.MAD.2021.111575>
- Booth, T. C., Nathan, M., Waldman, A. D., Quigley, A. M., Schapira, A. H., & Buscombe, J. (2015). The Role of Functional Dopamine-Transporter SPECT Imaging in Parkinsonian Syndromes, Part 1. *American Journal of Neuroradiology*, *36*(2), 229–235. <https://doi.org/10.3174/AJNR.A3970>
- Borchelt, D. R., Thinakaran, G., Eckman, C. B., Lee, M. K., Davenport, F., Ratovitsky, T., Prada, C. M., Kim, G., Seekins, S., Yager, D., Slunt, H. H., Wang, R., Seeger, M., Levey, A. I., Gandy, S. E., Copeland, N. G., Jenkins, N. A., Price, D. L., Younkin, S. G., & Sisodia, S. S. (1996). Familial Alzheimer's disease-linked presenilin I variants elevate  $a\beta_{1-42/1-40}$  ratio in vitro and in vivo. *Neuron*, *17*(5), 1005–1013. [https://doi.org/10.1016/S0896-6273\(00\)80230-5](https://doi.org/10.1016/S0896-6273(00)80230-5)
- Braak, H., & Braak, E. (1991). Neuropathological staging of Alzheimer-related changes. *Acta Neuropathologica*, *82*(4), 239–259. <https://doi.org/10.1007/BF00308809/METRICS>
- Braak, H., & Braak, E. (1996). Development of Alzheimer-related neurofibrillary changes in the neocortex inversely recapitulates cortical myelogenesis. *Acta Neuropathologica* *1996* *92:2*, *92*(2), 197–201. <https://doi.org/10.1007/S004010050508>
- Braak, H., & Del Tredici, K. (2015). The preclinical phase of the pathological process underlying sporadic Alzheimer's disease. *Brain*, *138*(10), 2814–2833. <https://doi.org/10.1093/BRAIN/AWV236>
- Braak, H., Del Tredici, K., Rüb, U., De Vos, R. A. I., Jansen Steur, E. N. H., & Braak, E. (2003). Staging of brain pathology related to sporadic Parkinson's disease. *Neurobiology of Aging*, *24*(2), 197–211. [https://doi.org/10.1016/S0197-4580\(02\)00065-9](https://doi.org/10.1016/S0197-4580(02)00065-9)
- Buchholz, S., & Zempel, H. (2024). The six brain-specific TAU isoforms and their role in Alzheimer's disease and related neurodegenerative dementia syndromes. *Alzheimer's & Dementia*. <https://doi.org/10.1002/ALZ.13784>
- Buckner, R. L., Snyder, A. Z., Shannon, B. J., LaRossa, G., Sachs, R., Fotenos, A. F., Sheline, Y. I., Klunk, W. E., Mathis, C. A., Morris, J. C., & Mintun, M. A. (2005). Molecular, Structural, and Functional Characterization of Alzheimer's Disease: Evidence for a Relationship between Default Activity, Amyloid, and Memory. *Journal of Neuroscience*, *25*(34), 7709–7717. <https://doi.org/10.1523/JNEUROSCI.2177-05.2005>
- Busche, M. A., Chen, X., Henning, H. A., Reichwald, J., Staufenbiel, M., Sakmann, B., & Konnerth, A. (2012). Critical role of soluble amyloid- $\beta$  for early hippocampal hyperactivity in a mouse model of Alzheimer's disease. *Proceedings of the National Academy of Sciences*, *109*(22), 8740–8745. <https://doi.org/10.1073/pnas.1206171109>

- Buzsáki, G. (1989). Two-stage model of memory trace formation: A role for “noisy” brain states. *Neuroscience*, *31*(3), 551–570. [https://doi.org/10.1016/0306-4522\(89\)90423-5](https://doi.org/10.1016/0306-4522(89)90423-5)
- Cannon, J., & Miller, P. (2016). Synaptic and intrinsic homeostasis cooperate to optimize single neuron response properties and tune integrator circuits. *Journal of Neurophysiology*, *116*(5), 2004–2022. <https://doi.org/10.1152/JN.00253.2016/ASSET/IMAGES/LARGE/Z9K0091637750013.JPEG>
- Chen, P., Shen, Z., Wang, Q., Zhang, B., Zhuang, Z., Lin, J., Shen, Y., Chen, Y., Dai, Z., & Wu, R. (2021). Reduced Cerebral Glucose Uptake in an Alzheimer’s Rat Model With Glucose-Weighted Chemical Exchange Saturation Transfer Imaging. *Frontiers in Aging Neuroscience*, *13*, 618690. <https://doi.org/10.3389/FNAGI.2021.618690/BIBTEX>
- Chen, Y., & Yu, Y. (2023). Tau and neuroinflammation in Alzheimer’s disease: interplay mechanisms and clinical translation. *Journal of Neuroinflammation* *2023* *20*:1, *20*(1), 1–21. <https://doi.org/10.1186/S12974-023-02853-3>
- Cho, H., Choi, J. Y., Hwang, M. S., Kim, Y. J., Lee, H. M., Lee, H. S., Lee, J. H., Ryu, Y. H., Lee, M. S., & Lyoo, C. H. (2016). In vivo cortical spreading pattern of tau and amyloid in the Alzheimer disease spectrum. *Annals of Neurology*, *80*(2), 247–258. <https://doi.org/10.1002/ANA.24711>
- Citron, M., Westaway, D., Xia, W., Carlson, G., Diehl, T., Levesque, G., Johnson-Wood, K., Lee, M., Seubert, P., Davis, A., Kholodenko, D., Motter, R., Sherrington, R., Perry, B., Yao, H., Strome, R., Lieberburg, I., Rommens, J., Kim, S., ... Selkoe, D. J. (1997). Mutant presenilins of Alzheimer’s disease increase production of 42-residue amyloid  $\beta$ -protein in both transfected cells and transgenic mice. *Nature Medicine* *1997* *3*:1, *3*(1), 67–72. <https://doi.org/10.1038/nm0197-67>
- Congdon, E. E., & Sigurdsson, E. M. (2018). Tau-targeting therapies for Alzheimer disease. *Nature Reviews Neurology* *2018* *14*:7, *14*(7), 399–415. <https://doi.org/10.1038/s41582-018-0013-z>
- Conte, M. I., Fuentes-Trillo, A., & Domínguez Conde, C. (2024). Opportunities and tradeoffs in single-cell transcriptomic technologies. *Trends in Genetics*, *40*(1), 83–93. <https://doi.org/10.1016/J.TIG.2023.10.003>
- Cortes-Canteli, M., & Iadecola, C. (2020). Alzheimer’s Disease and Vascular Aging: JACC Focus Seminar. *Journal of the American College of Cardiology*, *75*(8), 942–951. <https://doi.org/10.1016/J.JACC.2019.10.062>
- Cramb, K. M. L., Beccano-Kelly, D., Cragg, S. J., & Wade-Martins, R. (2023). Impaired dopamine release in Parkinson’s disease. *Brain*, *146*(8), 3117–3132. <https://doi.org/10.1093/BRAIN/AWAD064>
- Delaby, C., Hirtz, C., & Lehmann, S. (2023). Overview of the blood biomarkers in Alzheimer’s disease: Promises and challenges. *Revue Neurologique*, *179*(3), 161–172. <https://doi.org/https://doi.org/10.1016/j.neurol.2022.09.003>
- Dickerson, B. C., Salat, D. H., Greve, D. N., Chua, E. F., Rand-Giovannetti, E., Rentz, D. M., Bertram, L., Mullin, K., Tanzi, R. E., Blacker, D., Albert, M. S., & Sperling, R. A. (2005).

Increased hippocampal activation in mild cognitive impairment compared to normal aging and AD. *Neurology*, 65(3), 404–411. <https://doi.org/10.1212/01.WNL.0000171450.97464.49>

Dijkstra, A. A., Voorn, P., Berendse, H. W., Groenewegen, H. J., Rozemuller, A. J. M., & van de Berg, W. D. J. (2014). Stage-dependent nigral neuronal loss in incidental Lewy body and Parkinson's disease. *Movement Disorders*, 29(10), 1244–1251. <https://doi.org/10.1002/MDS.25952>

Dillman, A. A., Majounie, E., Ding, J., Gibbs, J. R., Hernandez, D., Arepalli, S., Traynor, B. J., Singleton, A. B., Galter, D., & Cookson, M. R. (2017). Transcriptomic profiling of the human brain reveals that altered synaptic gene expression is associated with chronological aging. *Scientific Reports 2017 7:1*, 7(1), 1–12. <https://doi.org/10.1038/s41598-017-17322-0>

Dodge, H. H., Woltjer, R., Kaye, J. A., Lyons, D. E., Mattek, N., Silbert, L. C., Zhu, J., Nelson, P. T., Fardo, D. W., Schmitt, F. A., Kryscio, R. J., Abner, E. L., Bennett, D. A., Schneider, J. A., Yu, L., Cairns, N. J., & Xiong, C. (2017). Risk of incident clinical diagnosis of Alzheimer's disease—type dementia attributable to pathology-confirmed vascular disease. *Alzheimer's & Dementia*, 13(6), 613–623. <https://doi.org/10.1016/J.JALZ.2016.11.003>

Donaghy, P. C., & McKeith, I. G. (2014). The clinical characteristics of dementia with Lewy bodies and a consideration of prodromal diagnosis. *Alzheimer's Research and Therapy*, 6(4), 1–12. <https://doi.org/10.1186/ALZRT274/FIGURES/7>

Dong-Chen, X., Yong, C., Yang, X., Chen-Yu, S., & Li-Hua, P. (2023). Signaling pathways in Parkinson's disease: molecular mechanisms and therapeutic interventions. *Signal Transduction and Targeted Therapy*, 8(1), 73. <https://doi.org/10.1038/s41392-023-01353-3>

Dukart, J., Kherif, F., Mueller, K., Adaszewski, S., Schroeter, M. L., Frackowiak, R. S. J., & Draganski, B. (2013). Generative FDG-PET and MRI Model of Aging and Disease Progression in Alzheimer's Disease. *PLOS Computational Biology*, 9(4), e1002987. <https://doi.org/10.1371/JOURNAL.PCBI.1002987>

Duncombe, J., Kitamura, A., Hase, Y., Ihara, M., Kalaria, R. N., & Horsburgh, K. (2017). Chronic cerebral hypoperfusion: a key mechanism leading to vascular cognitive impairment and dementia. Closing the translational gap between rodent models and human vascular cognitive impairment and dementia. *Clinical Science*, 131(19), 2451–2468. <https://doi.org/10.1042/CS20160727>

Ferrari, B. L., De Carvalho Campos Neto, G., Nucci, M. P., Mamani, J. B., Lacerda, S. S., Felício, A. C., Amaro, E., & Gamarra, L. F. (2019). The accuracy of hippocampal volumetry and glucose metabolism for the diagnosis of patients with suspected Alzheimer's disease, using automatic quantitative clinical tools. *Medicine (United States)*, 98(45). <https://doi.org/10.1097/MD.0000000000017824>

Fields, J. A., Ferman, T. J., Boeve, B. F., & Smith, G. E. (2011). Neuropsychological assessment of patients with dementing illness. *Nature Reviews Neurology 2011 7:12*, 7(12), 677–687. <https://doi.org/10.1038/nrneuro.2011.173>

- Fleisher, A. S., Chen, K., Quiroz, Y. T., Jakimovich, L. J., Gutierrez Gomez, M., Langois, C. M., Langbaum, J. B. S., Rontiva, A., Thiyyagura, P., Lee, W., Ayutyanont, N., Lopez, L., Moreno, S., Muñoz, C., Tirado, V., Acosta-Baena, N., Fagan, A. M., Giraldo, M., Garcia, G., ... Reiman, E. M. (2015). Associations Between Biomarkers and Age in the Presenilin 1 E280A Autosomal Dominant Alzheimer Disease Kindred: A Cross-sectional Study. *JAMA Neurology*, *72*(3), 316–324. <https://doi.org/10.1001/JAMANEUROL.2014.3314>
- Fonteijn, H. M., Modat, M., Clarkson, M. J., Barnes, J., Lehmann, M., Hobbs, N. Z., Schill, R. I., Tabrizi, S. J., Ourselin, S., Fox, N. C., & Alexander, D. C. (2012). An event-based model for disease progression and its application in familial Alzheimer's disease and Huntington's disease. *NeuroImage*, *60*(3), 1880–1889. <https://doi.org/10.1016/J.NEUROIMAGE.2012.01.062>
- Franceschi, C., Garagnani, P., Morsiani, C., Conte, M., Santoro, A., Grignolio, A., Monti, D., Capri, M., & Salvioli, S. (2018). The continuum of aging and age-related diseases: Common mechanisms but different rates. *Frontiers in Medicine*, *5*(MAR), 349810. <https://doi.org/10.3389/FMED.2018.00061/BIBTEX>
- Frisoni, G. B., Altomare, D., Thal, D. R., Ribaldi, F., van der Kant, R., Ossenkoppele, R., Blennow, K., Cummings, J., van Duijn, C., Nilsson, P. M., Dietrich, P. Y., Scheltens, P., & Dubois, B. (2021). The probabilistic model of Alzheimer disease: the amyloid hypothesis revised. *Nature Reviews Neuroscience* *2021* *23*:1, *23*(1), 53–66. <https://doi.org/10.1038/s41583-021-00533-w>
- Gammon, K. (2014). Neurodegenerative disease: Brain windfall. *Nature* *2014* *515*:7526, *515*(7526), 299–300. <https://doi.org/10.1038/nj7526-299a>
- Garbarino, S., Lorenzi, M., Oxtoby, N. P., Vinke, E. J., Marinescu, R. V., Eshaghi, A., Ikram, M. A., Niessen, W. J., Ciccarelli, O., Barkhof, F., Schott, J. M., Vernooij, M. W., & Alexander, D. C. (2019). Differences in topological progression profile among neurodegenerative diseases from imaging data. *ELife*, *8*. <https://doi.org/10.7554/ELIFE.49298>
- Garnier-Crussard, A., Cotton, F., Krolak-Salmon, P., & Chételat, G. (2023). White matter hyperintensities in Alzheimer's disease: Beyond vascular contribution. *Alzheimer's & Dementia*, *19*(8), 3738–3748. <https://doi.org/10.1002/ALZ.13057>
- Gerfen, C. R., Engber, T. M., Mahan, L. C., Susel, Z., Chase, T. N., Frederick J. Monsma, J., & Sibley, D. R. (1990). D1 and D2 Dopamine Receptor-regulated Gene Expression of Striatonigral and Striatopallidal Neurons. *Science*, *250*(4986), 1429–1432. <https://doi.org/10.1126/SCIENCE.2147780>
- Gerfen, C. R., & Surmeier, D. J. (2011). Modulation of striatal projection systems by dopamine. *Annual Review of Neuroscience*, *34*(Volume 34, 2011), 441–466. <https://doi.org/10.1146/ANNUREV-NEURO-061010-113641/CITE/REFWORKS>
- Gonzalez-Rodriguez, P., Zampese, E., & Surmeier, D. J. (2020). Selective neuronal vulnerability in Parkinson's disease. *Progress in Brain Research*, *252*, 61–89. <https://doi.org/10.1016/BS.PBR.2020.02.005>

- Gosche, K. M., Mortimer, J. A., Smith, C. D., Markesbery, W. R., & Snowdon, D. A. (2002). Hippocampal volume as an index of Alzheimer neuropathology: Findings from the Nun study. *Neurology*, *58*(10), 1476–1482. <https://doi.org/10.1212/WNL.58.10.1476>
- Grothe, M. J., Sepulcre, J., Gonzalez-Escamilla, G., Jelistratova, I., Schöll, M., Hansson, O., & Teipel, S. J. (2018). Molecular properties underlying regional vulnerability to Alzheimer's disease pathology. *Brain*, *141*(9), 2755–2771. <https://doi.org/10.1093/BRAIN/AWY189>
- Groves, W. C., Brandt, J., Steinberg, M., Warren, A., Rosenblatt, A., Baker, A., & Lyketsos, C. G. (2000). Vascular Dementia and Alzheimer's Disease: Is There a Difference? *The Journal of Neuropsychiatry and Clinical Neurosciences*, *12*(3), 305–315. <https://doi.org/10.1176/jnp.12.3.305>
- Guerreiro, R., & Bras, J. (2015). The age factor in Alzheimer's disease. *Genome Medicine*, *7*(1), 106. <https://doi.org/10.1186/s13073-015-0232-5>
- Guerrero, R., Schmidt-Richberg, A., Ledig, C., Tong, T., Wolz, R., & Rueckert, D. (2016). Instantiated mixed effects modeling of Alzheimer's disease markers. *NeuroImage*, *142*, 113–125. <https://doi.org/10.1016/J.NEUROIMAGE.2016.06.049>
- Hall, A. M., Throesch, B. T., Buckingham, S. C., Markwardt, S. J., Peng, Y., Wang, Q., Hoffman, D. A., & Roberson, E. D. (2015). Tau-Dependent Kv4.2 Depletion and Dendritic Hyperexcitability in a Mouse Model of Alzheimer's Disease. *Journal of Neuroscience*, *35*(15), 6221–6230. <https://doi.org/10.1523/JNEUROSCI.2552-14.2015>
- Halliday, G. M., & McCann, H. (2010). The progression of pathology in Parkinson's disease. *Annals of the New York Academy of Sciences*, *1184*(1), 188–195. <https://doi.org/10.1111/J.1749-6632.2009.05118.X>
- Hardy, J. A., & Higgins, G. A. (1992). Alzheimer's disease: The amyloid cascade hypothesis. *Science*, *256*(5054), 184–185. <https://doi.org/10.1126/SCIENCE.1566067/ASSET/74A758CB-215F-4A8E-9AAA-05F4D89DA62A/ASSETS/SCIENCE.1566067.FP.PNG>
- Hardy, J., & Allsop, D. (1991). Amyloid deposition as the central event in the aetiology of Alzheimer's disease. *Trends in Pharmacological Sciences*, *12*(C), 383–388. [https://doi.org/10.1016/0165-6147\(91\)90609-V](https://doi.org/10.1016/0165-6147(91)90609-V)
- Heemels, M. T. (2016). Neurodegenerative diseases. *Nature*, *539*(7628), 179. <https://doi.org/10.1038/539179A>
- Heo, J. Y., Nam, M. H., Yoon, H. H., Kim, J., Hwang, Y. J., Won, W., Woo, D. H., Lee, J. A., Park, H. J., Jo, S., Lee, M. J., Kim, S., Shim, J. E., Jang, D. P., Kim, K. I., Huh, S. H., Jeong, J. Y., Kowall, N. W., Lee, J., ... Lee, C. J. (2020). Aberrant Tonic Inhibition of Dopaminergic Neuronal Activity Causes Motor Symptoms in Animal Models of Parkinson's Disease. *Current Biology*, *30*(2), 276–291.e9. <https://doi.org/10.1016/J.CUB.2019.11.079>
- Höglinger, G. U., Adler, C. H., Berg, D., Klein, C., Outeiro, T. F., Poewe, W., Postuma, R., Stoessl, A. J., & Lang, A. E. (2024). A biological classification of Parkinson's disease: the SynNeurGe research diagnostic criteria. *The Lancet Neurology*, *23*(2), 191–204.

[https://doi.org/https://doi.org/10.1016/S1474-4422\(23\)00404-0](https://doi.org/https://doi.org/10.1016/S1474-4422(23)00404-0)

- Holdorff, B. (2002). Friedrich Heinrich Lewy (1885-1950) and his work. *Journal of the History of the Neurosciences*, *11*(1), 19–28. <https://doi.org/10.1076/JHIN.11.1.19.9106>
- Huang, C. J., Zhou, X., Yuan, X., Zhang, W., Li, M. X., You, M. Z., Zhu, X. Q., & Sun, Z. W. (2022). Contribution of Inflammation and Hypoperfusion to White Matter Hyperintensities-Related Cognitive Impairment. *Frontiers in Neurology*, *12*. <https://doi.org/10.3389/FNEUR.2021.786840/FULL>
- Huang, J., van Zijl, P. C. M., Han, X., Dong, C. M., Cheng, G. W. Y., Tse, K.-H., Knutsson, L., Chen, L., Lai, J. H. C., Wu, E. X., Xu, J., & Chan, K. W. Y. (2024). Altered d-glucose in brain parenchyma and cerebrospinal fluid of early Alzheimer's disease detected by dynamic glucose-enhanced MRI. *Science Advances*, *6*(20), eaba3884. <https://doi.org/10.1126/sciadv.aba3884>
- Inoue, Y., Shue, F., Bu, G., & Kanekiyo, T. (2023). Pathophysiology and probable etiology of cerebral small vessel disease in vascular dementia and Alzheimer's disease. *Molecular Neurodegeneration* *2023 18:1*, *18*(1), 1–22. <https://doi.org/10.1186/S13024-023-00640-5>
- Iturria-Medina, Y., Carbonell, F. M., Sotero, R. C., Chouinard-Decorte, F., Evans, A. C., Y, I.-M., FM, C., RC, S., F, C.-D., AC, E., Iturria-Medina, Y., Carbonell, F. M., Sotero, R. C., Chouinard-Decorte, F., & Evans, A. C. (2017). Multifactorial causal model of brain (dis)organization and therapeutic intervention: Application to Alzheimer's disease. *NeuroImage*, *152*, 60–77. <https://doi.org/10.1016/j.neuroimage.2017.02.058>
- Iturria-Medina, Y., Khan, A. F., Adewale, Q., Shirazi, A. H., & Initiative, the A. D. N. (2020). Blood and brain gene expression trajectories mirror neuropathology and clinical deterioration in neurodegeneration. *Brain*, *143*(2), 661–673. <https://doi.org/10.1093/BRAIN/AWZ400>
- Iturria-Medina, Y., Sotero, R. C., Toussaint, P. J., & Evans, A. C. (2014). Epidemic Spreading Model to Characterize Misfolded Proteins Propagation in Aging and Associated Neurodegenerative Disorders. *PLOS Computational Biology*, *10*(11), e1003956. <https://doi.org/10.1371/JOURNAL.PCBI.1003956>
- Iturria-Medina, Y., Sotero, R. C., Toussaint, P. J., Mateos-Pérez, J. M., Evans, A. C., Weiner, M. M. W., Aisen, P., Petersen, R., Jack, C. R., Jagust, W., Trojanowki, J. Q., Toga, A. W., Beckett, L., Green, R. C., Saykin, A. J., Morris, J., Shaw, L. M., Khachaturian, Z., Sorensen, G., ... Furst, A. J. (2016). Early role of vascular dysregulation on late-onset Alzheimer's disease based on multifactorial data-driven analysis. *Nature Communications*, *7*(1), 1–14. <https://doi.org/10.1038/ncomms11934>
- Jack, C. R., Barkhof, F., Bernstein, M. A., Cantillon, M., Cole, P. E., Decarli, C., Dubois, B., Duchesne, S., Fox, N. C., Frisoni, G. B., Hampel, H., Hill, D. L. G., Johnson, K., Mangin, J. F., Scheltens, P., Schwarz, A. J., Sperling, R., Suhy, J., Thompson, P. M., ... Foster, N. L. (2011). Steps to standardization and validation of hippocampal volumetry as a biomarker in clinical trials and diagnostic criterion for Alzheimer's disease. *Alzheimer's & Dementia*, *7*(4), 474-485.e4. <https://doi.org/10.1016/J.JALZ.2011.04.007>

- Jack, C. R., Dickson, D. W., Parisi, J. E., Xu, Y. C., Cha, R. H., O'Brien, P. C., Edland, S. D., Smith, G. E., Boeve, B. F., Tangalos, E. G., Kokmen, E., & Petersen, R. C. (2002). Antemortem MRI findings correlate with hippocampal neuropathology in typical aging and dementia. *Neurology*, *58*(5), 750–757. <https://doi.org/10.1212/WNL.58.5.750>
- Jack, C. R., Knopman, D. S., Jagust, W. J., Shaw, L. M., Aisen, P. S., Weiner, M. W., Petersen, R. C., & Trojanowski, J. Q. (2010). Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *The Lancet Neurology*, *9*(1), 119–128. [https://doi.org/10.1016/S1474-4422\(09\)70299-6](https://doi.org/10.1016/S1474-4422(09)70299-6)
- Jack, C. R., Vemuri, P., Wiste, H. J., Weigand, S. D., Lesnick, T. G., Lowe, V., Kantarci, K., Bernstein, M. A., Senjem, M. L., Gunter, J. L., Boeve, B. F., Trojanowski, J. Q., Shaw, L. M., Aisen, P. S., Weiner, M. W., Petersen, R. C., & Knopman, D. S. (2012). Shapes of the Trajectories of 5 Major Biomarkers of Alzheimer Disease. *Archives of Neurology*, *69*(7), 856–867. <https://doi.org/10.1001/ARCHNEUROL.2011.3405>
- Jack, C. R., Wiste, H. J., Schwarz, C. G., Lowe, V. J., Senjem, M. L., Vemuri, P., Weigand, S. D., Therneau, T. M., Knopman, D. S., Gunter, J. L., Jones, D. T., Graff-Radford, J., Kantarci, K., Roberts, R. O., Mielke, M. M., MacHulda, M. M., & Petersen, R. C. (2018). Longitudinal tau PET in ageing and Alzheimer's disease. *Brain : A Journal of Neurology*, *141*(5), 1517–1528. <https://doi.org/10.1093/BRAIN/AWY059>
- Jacobs, H. I. L., Radua, J., Lückmann, H. C., & Sack, A. T. (2013). Meta-analysis of functional network alterations in Alzheimer's disease: Toward a network biomarker. *Neuroscience & Biobehavioral Reviews*, *37*(5), 753–765. <https://doi.org/10.1016/J.NEUBIOREV.2013.03.009>
- Jasinska, A. J., Service, S., Choi, O., DeYoung, J., Grujic, O., Kong, S., Jorgensen, M. J., Bailey, J., Breidenthal, S., Fairbanks, L. A., Woods, R. P., Jentsch, J. D., & Freimer, N. B. (2009). Identification of brain transcriptional variation reproduced in peripheral blood: an approach for mapping brain expression traits. *Human Molecular Genetics*, *18*(22), 4415–4427. <https://doi.org/10.1093/hmg/ddp397>
- Jellinger, K. A. (2010). Basic mechanisms of neurodegeneration: a critical update. *Journal of Cellular and Molecular Medicine*, *14*(3), 457. <https://doi.org/10.1111/J.1582-4934.2010.01010.X>
- Jellinger, K. A., & Attems, J. (2010). Is there pure vascular dementia in old age? *Journal of the Neurological Sciences*, *299*(1), 150–154. <https://doi.org/https://doi.org/10.1016/j.jns.2010.08.038>
- Johnson, E. C. B., Dammer, E. B., Duong, D. M., Ping, L., Zhou, M., Yin, L., Higginbotham, L. A., Guajardo, A., White, B., Troncoso, J. C., Thambisetty, M., Montine, T. J., Lee, E. B., Trojanowski, J. Q., Beach, T. G., Reiman, E. M., Haroutunian, V., Wang, M., Schadt, E., ... Seyfried, N. T. (2020). Large-scale proteomic analysis of Alzheimer's disease brain and cerebrospinal fluid reveals early changes in energy metabolism associated with microglia and astrocyte activation. *Nature Medicine* *2020* *26*:5, *26*(5), 769–780. <https://doi.org/10.1038/s41591-020-0815-6>
- Joie, R. La, Visani, A. V., Baker, S. L., Brown, J. A., Bourakova, V., Cha, J., Chaudhary, K., Edwards, L., Iaccarino, L., Janabi, M., Lesman-Segev, O. H., Miller, Z. A., Perry, D. C.,

- O’Neil, J. P., Pham, J., Rojas, J. C., Rosen, H. J., Seeley, W. W., Tsai, R. M., ... Rabinovici, G. D. (2020). Prospective longitudinal atrophy in Alzheimer’s disease correlates with the intensity and topography of baseline tau-PET. *Science Translational Medicine*, *12*(524), 5732. [https://doi.org/10.1126/SCITRANSLMED.AAU5732/SUPPL\\_FILE/AAU5732\\_SM.PDF](https://doi.org/10.1126/SCITRANSLMED.AAU5732/SUPPL_FILE/AAU5732_SM.PDF)
- Josephs, K. A., Murray, M. E., Whitwell, J. L., Parisi, J. E., Petrucelli, L., Jack, C. R., Petersen, R. C., & Dickson, D. W. (2014). Staging TDP-43 pathology in Alzheimer’s disease. *Acta Neuropathologica*, *127*(3), 441–450. <https://doi.org/10.1007/S00401-013-1211-9/FIGURES/3>
- Juárez Olgún, H., Calderón Guzmán, D., Hernández García, E., & Barragán Mejía, G. (2016). The Role of Dopamine and Its Dysfunction as a Consequence of Oxidative Stress. *Oxidative Medicine and Cellular Longevity*, 2016. <https://doi.org/10.1155/2016/9730467>
- Jucker, M., & Walker, L. C. (2013). Self-propagation of pathogenic protein aggregates in neurodegenerative diseases. *Nature* 2013 *501*:7465, *501*(7465), 45–51. <https://doi.org/10.1038/nature12481>
- Kampmann, M. (2024). Molecular and cellular mechanisms of selective vulnerability in neurodegenerative diseases. *Nature Reviews Neuroscience* 2024 *25*:5, *25*(5), 351–371. <https://doi.org/10.1038/s41583-024-00806-0>
- Kesidou, E., Theotokis, P., Damianidou, O., Boziki, M., Konstantinidou, N., Taloumtzis, C., Sintila, S. A., Grigoriadis, P., Evangelopoulos, M. E., Bakirtzis, C., & Simeonidou, C. (2023). CNS Ageing in Health and Neurodegenerative Disorders. *Journal of Clinical Medicine* 2023, Vol. 12, Page 2255, *12*(6), 2255. <https://doi.org/10.3390/JCM12062255>
- Kim, C. K., Lee, Y. R., Ong, L., Gold, M., Kalali, A., & Sarkar, J. (2022). Alzheimer’s Disease: Key Insights from Two Decades of Clinical Trial Failures. *Journal of Alzheimer’s Disease*, *87*(1), 83. <https://doi.org/10.3233/JAD-215699>
- Köpke, E., Tung, Y. C., Shaikh, S., Alonso, A. C., Iqbal, K., & Grundke-Iqbal, I. (1993). Microtubule-associated protein tau. Abnormal phosphorylation of a non-paired helical filament pool in Alzheimer disease. *The Journal of Biological Chemistry*, *268*(32), 24374–24384.
- Lee, H., Lee, J. J., Park, N. Y., Dubey, S. K., Kim, T., Ruan, K., Lim, S. Bin, Park, S. H., Ha, S., Kovlyagina, I., Kim, K. tai, Kim, S., Oh, Y., Kim, H., Kang, S. U., Song, M. R., Lloyd, T. E., Maragakis, N. J., Hong, Y. Bin, ... Lee, G. (2021). Multi-omic analysis of selectively vulnerable motor neuron subtypes implicates altered lipid metabolism in ALS. *Nature Neuroscience* 2021 *24*:12, *24*(12), 1673–1685. <https://doi.org/10.1038/s41593-021-00944-z>
- Lee, W. J., Brown, J. A., Kim, H. R., La Joie, R., Cho, H., Lyoo, C. H., Rabinovici, G. D., Seong, J. K., & Seeley, W. W. (2022). Regional A $\beta$ -tau interactions promote onset and acceleration of Alzheimer’s disease tau spreading. *Neuron*, *110*(12), 1932–1943.e5. <https://doi.org/10.1016/j.neuron.2022.03.034>
- Leng, K., Li, E., Eser, R., Piergies, A., Sit, R., Tan, M., Neff, N., Li, S. H., Rodriguez, R. D.,

- Suemoto, C. K., Leite, R. E. P., Ehrenberg, A. J., Pasqualucci, C. A., Seeley, W. W., Spina, S., Heinsen, H., Grinberg, L. T., & Kampmann, M. (2021). Molecular characterization of selectively vulnerable neurons in Alzheimer's disease. *Nature Neuroscience*, *24*(2), 276–287. <https://doi.org/10.1038/s41593-020-00764-7>
- Leuzy, A., Mattsson-Carlsson, N., Palmqvist, S., Janelidze, S., Dage, J. L., & Hansson, O. (2022). Blood-based biomarkers for Alzheimer's disease. *EMBO Molecular Medicine*, *14*(1), e14408. <https://doi.org/10.15252/emmm.202114408>
- Liu, X., Wang, S., Zhang, X., Wang, Z., Tian, X., & He, Y. (2014). Abnormal amplitude of low-frequency fluctuations of intrinsic brain activity in Alzheimer's disease. *Journal of Alzheimer's Disease : JAD*, *40*(2), 387–397. <https://doi.org/10.3233/JAD-131322>
- Maiti, P., Manna, J., & Dunbar, G. L. (2017). Current understanding of the molecular mechanisms in Parkinson's disease: Targets for potential treatments. *Translational Neurodegeneration*, *6*(1), 28. <https://doi.org/10.1186/s40035-017-0099-z>
- Marsh, B., & Brelloch, R. (2020). Single nuclei RNA-seq of mouse placental labyrinth development. *ELife*, *9*, 1–27. <https://doi.org/10.7554/ELIFE.60266>
- Martí-Juan, G., Lorenzi, M., & Piella, G. (2023). MC-RVAE: Multi-channel recurrent variational autoencoder for multimodal Alzheimer's disease progression modelling. *NeuroImage*, *268*, 119892. <https://doi.org/10.1016/J.NEUROIMAGE.2023.119892>
- McCarthy, J. (2022). *Data-driven MRI modelling for the characterization, staging, and subtyping of frontotemporal dementia* [McGill University]. <https://escholarship.mcgill.ca/concern/theses/kw52jd91v>
- McDonald, C. R., Gharapetian, L., McEvoy, L. K., Fennema-Notestine, C., Hagler, D. J., Holland, D., & Dale, A. M. (2012). Relationship between regional atrophy rates and cognitive decline in mild cognitive impairment. *Neurobiology of Aging*, *33*(2), 242–253. <https://doi.org/10.1016/J.NEUROBIOLAGING.2010.03.015>
- McGregor, M. M., & Nelson, A. B. (2019). Circuit Mechanisms of Parkinson's Disease. *Neuron*, *101*(6), 1042–1056. <https://doi.org/10.1016/j.neuron.2019.03.004>
- McKenzie, A. T., Wang, M., Hauberg, M. E., Fullard, J. F., Kozlenkov, A., Keenan, A., Hurd, Y. L., Dracheva, S., Casaccia, P., Roussos, P., & Zhang, B. (2018). Brain Cell Type Specific Gene Expression and Co-expression Network Architectures. *Scientific Reports*, *8*(1), 8868. <https://doi.org/10.1038/s41598-018-27293-5>
- McNay, E. C., Ong, C. T., McCrimmon, R. J., Cresswell, J., Bogan, J. S., & Sherwin, R. S. (2010). Hippocampal memory processes are modulated by insulin and high-fat-induced insulin resistance. *Neurobiology of Learning and Memory*, *93*(4), 546–553. <https://doi.org/10.1016/J.NLM.2010.02.002>
- Mi, Y. J., Won, Y. L., Ho, Y. K., & Eun, J. C. (2007). The effects of L-3,4-dihydroxyphenylalanine and dopamine agonists on dopamine neurons in the progressive hemiparkinsonian rat models. *Neurological Research*, *29*(3), 289–295. <https://doi.org/10.1179/174313206X153996>
- Mihaescu, A. S., Valli, M., Uribe, C., Diez-Cirarda, M., Masellis, M., Graff-Guerrero, A., &

- Strafella, A. P. (2022). Beta amyloid deposition and cognitive decline in Parkinson's disease: a study of the PPMI cohort. *Molecular Brain*, 15(1), 1–13. <https://doi.org/10.1186/S13041-022-00964-1/TABLES/4>
- Morris, G. P., Clark, I. A., & Vissel, B. (2014). Inconsistencies and Controversies Surrounding the Amyloid Hypothesis of Alzheimer's Disease. *Acta Neuropathologica Communications*, 2(1), 1–21. <https://doi.org/10.1186/S40478-014-0135-5/FIGURES/3>
- Mostafavi, S., Gaiteri, C., Sullivan, S. E., White, C. C., Tasaki, S., Xu, J., Taga, M., Klein, H. U., Patrick, E., Komashko, V., McCabe, C., Smith, R., Bradshaw, E. M., Root, D. E., Regev, A., Yu, L., Chibnik, L. B., Schneider, J. A., Young-Pearse, T. L., ... De Jager, P. L. (2018). A molecular network of the aging human brain provides insights into the pathology and cognitive decline of Alzheimer's disease. *Nature Neuroscience*, 21(6), 811–819. <https://doi.org/10.1038/s41593-018-0154-9>
- Muralidar, S., Ambi, S. V., Sekaran, S., Thirumalai, D., & Palaniappan, B. (2020). Role of tau protein in Alzheimer's disease: The prime pathological player. *International Journal of Biological Macromolecules*, 163, 1599–1617. <https://doi.org/10.1016/J.IJBIOMAC.2020.07.327>
- Nakanishi, T., Sano, R., Kuwatsuru, S., Aoki, N., Hattori, K., Kamagata, Y., Motoi, O., Abe, K., Shimoji, M., Hori, A., Kamagata, K., Motoi, Y., Abe, O., Shimoji, K., Hori, M., Nakanishi, A., Sano, T., Kuwatsuru, R., Aoki, S., & Hattori, N. (2012). White Matter Alteration of the Cingulum in Parkinson Disease with and without Dementia: Evaluation by Diffusion Tensor Tract-Specific Analysis. *American Journal of Neuroradiology*, 33(5), 890–895. <https://doi.org/10.3174/AJNR.A2860>
- Nazmuddin, M., van Dalen, J. W., Borra, R. J. H., Stormezand, G. N., van der Horn, H. J., van der Zee, S., Boertien, J., & van Laar, T. (2021). Postural and gait symptoms in de novo Parkinson's disease patients correlate with cholinergic white matter pathology. *Parkinsonism and Related Disorders*, 93, 43–49. <https://doi.org/10.1016/j.parkreldis.2021.11.010>
- Newbold, D. J., Laumann, T. O., Hoyt, C. R., Hampton, J. M., Montez, D. F., Raut, R. V., Ortega, M., Mitra, A., Nielsen, A. N., Miller, D. B., Adeyemo, B., Nguyen, A. L., Scheidter, K. M., Tanenbaum, A. B., Van, A. N., Marek, S., Schlaggar, B. L., Carter, A. R., Greene, D. J., ... Dosenbach, N. U. F. (2020). Plasticity and Spontaneous Activity Pulses in Disused Human Brain Circuits. *Neuron*, 107(3), 580-589.e6. <https://doi.org/https://doi.org/10.1016/j.neuron.2020.05.007>
- Newman, A. M., Steen, C. B., Liu, C. L., Gentles, A. J., Chaudhuri, A. A., Scherer, F., Khodadoust, M. S., Esfahani, M. S., Luca, B. A., Steiner, D., Diehn, M., & Alizadeh, A. A. (2019). Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nature Biotechnology*, 37(7), 773–782. <https://doi.org/10.1038/s41587-019-0114-2>
- O'Brien, J. L., O'Keefe, K. M., Laviolette, P. S., Deluca, A. N., Blacker, D., Dickerson, B. C., & Sperling, R. A. (2010). Longitudinal fMRI in elderly reveals loss of hippocampal activation with clinical decline. *Neurology*, 74(24), 1969–1976. <https://doi.org/10.1212/WNL.0B013E3181E3966E/ASSET/732ABEE6-B1DF-47E9-BE1E-56AC231DC790/ASSETS/GRAPHIC/PAP.GIF>

- Obeso, J. A., & Lanciego, J. L. (2011). Past, present, and future of the pathophysiological model of the basal ganglia. *Frontiers in Neuroanatomy*, 5(JULY), 11718. <https://doi.org/10.3389/FNANA.2011.00039/BIBTEX>
- Ólafsdóttir, H. F., Bush, D., & Barry, C. (2018). The Role of Hippocampal Replay in Memory and Planning. *Current Biology*, 28(1), R37–R50. <https://doi.org/10.1016/J.CUB.2017.10.073>
- Olanow, C. W., & Brundin, P. (2013). Parkinson's Disease and Alpha Synuclein: Is Parkinson's Disease a Prion-Like Disorder? *Movement Disorders*, 28(1), 31–40. <https://doi.org/10.1002/MDS.25373>
- Pak, V., Adewale, Q., Bzdok, D., Dadar, M., Zeighami, Y., & Iturria-Medina, Y. (2023). *Distinctive Whole-brain Cell Types Predict Tissue Damage Patterns in Thirteen Neurodegenerative Conditions*. Cold Spring Harbor Laboratory. <https://doi.org/10.1101/2023.06.08.544227>
- Pang, Y., Kukull, W., Sano, M., Albin, R. L., Shen, C., Zhou, J., & Dodge, H. H. (2023). Predicting Progression from Normal to MCI and from MCI to AD Using Clinical Variables in the National Alzheimer's Coordinating Center Uniform Data Set Version 3: Application of Machine Learning Models and a Probability Calculator. *The Journal of Prevention of Alzheimer's Disease*, 10(2), 301–313. <https://doi.org/10.14283/jpad.2023.10>
- Papadimitriou, C. H., Vempala, S. S., Mitropolsky, D., Collins, M., & Maass, W. (2020). Brain computation by assemblies of neurons. *Proceedings of the National Academy of Sciences*, 117(25), 14464–14472. <https://doi.org/10.1073/pnas.2001893117>
- Parkinson, J. (2002). An essay on the shaking palsy. 1817. *The Journal of Neuropsychiatry and Clinical Neurosciences*, 14(2). <https://doi.org/10.1176/JNP.14.2.223>
- Pasanen, P., Myllykangas, L., Siitonen, M., Raunio, A., Kaakkola, S., Lyytinen, J., Tienari, P. J., Pöyhönen, M., & Paetau, A. (2014). A novel  $\alpha$ -synuclein mutation A53E associated with atypical multiple system atrophy and Parkinson's disease-type pathology. *Neurobiology of Aging*, 35(9), 2180.e1-2180.e5. <https://doi.org/https://doi.org/10.1016/j.neurobiolaging.2014.03.024>
- Patriat, R., Pisharady, P. K., Amundsen-Huffmaster, S., Linn-Evans, M., Howell, M., Chung, J. W., Petrucci, M. N., Videnovic, A., Holker, E., De Kam, J., Tuite, P., Lenglet, C., Harel, N., & MacKinnon, C. D. (2022). White matter microstructure in Parkinson's disease with and without elevated rapid eye movement sleep muscle tone. *Brain Communications*, 4(2). <https://doi.org/10.1093/BRAINCOMMS/FCAC027>
- Pearson-Leary, J., & McNay, E. C. (2016). Novel Roles for the Insulin-Regulated Glucose Transporter-4 in Hippocampally Dependent Memory. *Journal of Neuroscience*, 36(47), 11851–11864. <https://doi.org/10.1523/JNEUROSCI.1700-16.2016>
- Philipe de Souza Ferreira, L., André da Silva, R., Marques Mesquita da Costa, M., Moraes de Paiva Roda, V., Vizcaino, S., Janisset, N. R. L. L., Ramos Vieira, R., Marcos Sanches, J., Maria Soares Junior, J., & de Jesus Simões, M. (2022). Sex differences in Parkinson's Disease: An emerging health question. *Clinics*, 77.

<https://doi.org/10.1016/J.CLINSP.2022.100121>

- Pinares-Garcia, P., Stratikopoulos, M., Zagato, A., Loke, H., & Lee, J. (2018). Sex: A Significant Risk Factor for Neurodevelopmental and Neurodegenerative Disorders. *Brain Sciences*, 8(8), E154–E154. <https://doi.org/10.3390/BRAINSCI8080154>
- Planche, V., Manjon, J. V, Mansencal, B., Lanuza, E., Tourdias, T., Catheline, G., & Coupé, P. (2022). Structural progression of Alzheimer’s disease over decades: the MRI staging scheme. *Brain Communications*, 4(3), fcac109. <https://doi.org/10.1093/braincomms/fcac109>
- Radulescu, C. I., Doostdar, N., Zabouri, N., Melgosa-Ecenarro, L., Wang, X., Sadeh, S., Pavlidi, P., Airey, J., Kopanitsa, M., Clopath, C., & Barnes, S. J. (2023). Age-related dysregulation of homeostatic control in neuronal microcircuits. *Nature Neuroscience*, 26(12), 2158–2170. <https://doi.org/10.1038/s41593-023-01451-z>
- Raj, A. (2021). Graph Models of Pathology Spread in Alzheimer’s Disease: An Alternative to Conventional Graph Theoretic Analysis. <https://Home.Liebertpub.Com/Brain>, 11(10), 799–814. <https://doi.org/10.1089/BRAIN.2020.0905>
- Raj, A., Kuceyeski, A., & Weiner, M. (2012). A Network Diffusion Model of Disease Progression in Dementia. *Neuron*, 73(6), 1204–1215. <https://doi.org/10.1016/j.neuron.2011.12.040>
- Ramesh, S., Arachchige, A. S. P. M., Ramesh, S., & Arachchige, A. S. P. M. (2023). Depletion of dopamine in Parkinson’s disease and relevant therapeutic options: A review of the literature. *AIMS Neuroscience* 2023 3:200, 10(3), 200–231. <https://doi.org/10.3934/NEUROSCIENCE.2023017>
- Rektor, I., Svátková, A., Vojtíšek, L., Zikmundová, I., Vaníček, J., Király, A., & Szabó, N. (2018). White matter alterations in Parkinson’s disease with normal cognition precede grey matter atrophy. *PLOS ONE*, 13(1), e0187939. <https://doi.org/10.1371/JOURNAL.PONE.0187939>
- Rink, C., & Khanna, S. (2011). Significance of Brain Tissue Oxygenation and the Arachidonic Acid Cascade in Stroke. *Antioxidants & Redox Signaling*, 14(10), 1889. <https://doi.org/10.1089/ARS.2010.3474>
- Rittman, T., Rubinov, M., Vértes, P. E., Patel, A. X., Ginestet, C. E., Ghosh, B. C. P., Barker, R. A., Spillantini, M. G., Bullmore, E. T., & Rowe, J. B. (2016). Regional expression of the MAPT gene is associated with loss of hubs in brain networks and cognitive impairment in Parkinson disease and progressive supranuclear palsy. *Neurobiology of Aging*, 48, 153–160. <https://doi.org/10.1016/J.NEUROBIOLAGING.2016.09.001>
- Robinson, J. L., Lee, E. B., Xie, S. X., Rennert, L., Suh, E., Bredenberg, C., Caswell, C., Van Deerlin, V. M., Yan, N., Yousef, A., Hurtig, H. I., Siderowf, A., Grossman, M., McMillan, C. T., Miller, B., Duda, J. E., Irwin, D. J., Wolk, D., Elman, L., ... Trojanowski, J. Q. (2018). Neurodegenerative disease concomitant proteinopathies are prevalent, age-related and APOE4-associated. *Brain*, 141(7), 2181–2193. <https://doi.org/10.1093/BRAIN/AWY146>
- Rodrigue, K. M., Kennedy, K. M., Devous, M. D., Rieck, J. R., Hebrank, A. C., Diaz-Arrastia,

- R., Mathews, D., & Park, D. C. (2012).  $\beta$ -amyloid burden in healthy aging: Regional distribution and cognitive consequences. *Neurology*, 78(6), 387–395. [https://doi.org/10.1212/WNL.0B013E318245D295/SUPPL\\_FILE/WNL203400-399147\\_FIGURE\\_E-1.PDF](https://doi.org/10.1212/WNL.0B013E318245D295/SUPPL_FILE/WNL203400-399147_FIGURE_E-1.PDF)
- Santiago, C. P., Gimmen, M. Y., Lu, Y., McNally, M. M., Duncan, L. H., Creamer, T. J., Orzolek, L. D., Blackshaw, S., & Singh, M. S. (2023). Comparative Analysis of Single-cell and Single-nucleus RNA-sequencing in a Rabbit Model of Retinal Detachment-related Proliferative Vitreoretinopathy. *Ophthalmology Science*, 3(4), 100335. <https://doi.org/10.1016/J.XOPS.2023.100335>
- Schapira, A. H. V., Chaudhuri, K. R., & Jenner, P. (2017). Non-motor features of Parkinson disease. *Nature Reviews Neuroscience* 2017 18:7, 18(7), 435–450. <https://doi.org/10.1038/nrn.2017.62>
- Scheuner, D., Eckman, C., Jensen, M., Song, X., Citron, M., Suzuki, N., Bird, T. D., Hardy, J., Hutton, M., Kukull, W., Larson, E., Levy-Lahad, E., Viitanen, M., Peskind, E., Poorkaj, P., Schellenberg, G., Tanzi, R., Wasco, W., Lannfelt, L., ... Younkin, S. (1996). Secreted amyloid  $\beta$ -protein similar to that in the senile plaques of Alzheimer's disease is increased in vivo by the presenilin 1 and 2 and APP mutations linked to familial Alzheimer's disease. *Nature Medicine* 1996 2:8, 2(8), 864–870. <https://doi.org/10.1038/nm0896-864>
- Schulze, M., Sommer, A., Plötz, S., Farrell, M., Winner, B., Grosch, J., Winkler, J., & Riemenschneider, M. J. (2018). Sporadic Parkinson's disease derived neuronal cells show disease-specific mRNA and small RNA signatures with abundant deregulation of piRNAs. *Acta Neuropathologica Communications*, 6(1), 58. <https://doi.org/10.1186/s40478-018-0561-x>
- Sedzikowska, A., & Szablewski, L. (2021). Insulin and insulin resistance in alzheimer's disease. *International Journal of Molecular Sciences*, 22(18), 9987. <https://doi.org/10.3390/ijms22189987>
- Sepulcre, J., Grothe, M. J., d'Oleire Uquillas, F., Ortiz-Terán, L., Diez, I., Yang, H.-S., Jacobs, H. I. L., Hanseeuw, B. J., Li, Q., El-Fakhri, G., Sperling, R. A., & Johnson, K. A. (2018). Neurogenetic contributions to amyloid beta and tau spreading in the human cortex. *Nature Medicine*, 24(12), 1910–1918. <https://doi.org/10.1038/s41591-018-0206-4>
- Simuni, T., Chahine, L. M., Poston, K., Brumm, M., Buracchio, T., Campbell, M., Chowdhury, S., Coffey, C., Concha-Marambio, L., Dam, T., DiBiaso, P., Foroud, T., Frasier, M., Gochanour, C., Jennings, D., Kiebertz, K., Kopil, C. M., Merchant, K., Mollenhauer, B., ... Marek, K. (2024). A biological definition of neuronal  $\alpha$ -synuclein disease: towards an integrated staging system for research. *The Lancet Neurology*, 23(2), 178–190. [https://doi.org/https://doi.org/10.1016/S1474-4422\(23\)00405-2](https://doi.org/https://doi.org/10.1016/S1474-4422(23)00405-2)
- Singh, P. P., Demmitt, B. A., Nath, R. D., & Brunet, A. (2019). The Genetics of Aging: A Vertebrate Perspective. *Cell*, 177(1), 200–220. <https://doi.org/10.1016/J.CELL.2019.02.038>
- Šišková, Z., Justus, D., Kaneko, H., Friedrichs, D., Henneberg, N., Beutel, T., Pitsch, J., Schoch, S., Becker, A., vonderKammer, H., & Remy, S. (2014). Dendritic structural degeneration is functionally linked to cellular hyperexcitability in a mouse model of

- alzheimer's disease. *Neuron*, 84(5), 1023–1033.  
<https://doi.org/10.1016/j.neuron.2014.10.024>
- Stelzmann, R. A., Norman Schnitzlein, H., & Reed Murtagh, F. (1995). An english translation of alzheimer's 1907 paper, "über eine eigenartige erkankung der hirnrinde." *Clinical Anatomy*, 8(6), 429–431. <https://doi.org/10.1002/CA.980080612>
- Sullivan, P. F., Fan, C., & Perou, C. M. (2006). Evaluating the comparability of gene expression in blood and brain. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 141B(3), 261–268.  
<https://doi.org/https://doi.org/10.1002/ajmg.b.30272>
- Sweeney, M. D., Montagne, A., Sagare, A. P., Nation, D. A., Schneider, L. S., Chui, H. C., Harrington, M. G., Pa, J., Law, M., Wang, D. J. J., Jacobs, R. E., Doubal, F. N., Ramirez, J., Black, S. E., Nedergaard, M., Benveniste, H., Dichgans, M., Iadecola, C., Love, S., ... Zlokovic, B. V. (2019). Vascular dysfunction—The disregarded partner of Alzheimer's disease. *Alzheimer's & Dementia*, 15(1), 158–167.  
<https://doi.org/10.1016/J.JALZ.2018.07.222>
- Tanaka, T., Biancotto, A., Moaddel, R., Moore, A. Z., Gonzalez-Freire, M., Aon, M. A., Candia, J., Zhang, P., Cheung, F., Fantoni, G., Semba, R. D., & Ferrucci, L. (2018). Plasma proteomic signature of age in healthy humans. *Aging Cell*, 17(5), e12799.  
<https://doi.org/10.1111/ACEL.12799>
- Tandon, R., Kirkpatrick, A., & Mitchell, C. S. (2023). sEBM: Scaling Event Based Models to Predict Disease Progression via Implicit Biomarker Selection and Clustering. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 13939 LNCS, 208–221.  
[https://doi.org/10.1007/978-3-031-34048-2\\_17/FIGURES/4](https://doi.org/10.1007/978-3-031-34048-2_17/FIGURES/4)
- Taylor, J. P., Hardy, J., & Fischbeck, K. H. (2002). Toxic Proteins in Neurodegenerative Disease. *Science*, 296(5575), 1991–1995. <https://doi.org/10.1126/SCIENCE.1067122>
- Terwel, D., Dewachter, I., & Van Leuven, F. (2002). Axonal transport, tau protein, and neurodegeneration in Alzheimer's disease. *NeuroMolecular Medicine*, 2(2), 151–165.  
<https://doi.org/10.1385/NMM:2:2:151/METRICS>
- Thal, D. R., Rüb, U., Orantes, M., & Braak, H. (2002). Phases of A $\beta$ -deposition in the human brain and its relevance for the development of AD. *Neurology*, 58(12), 1791–1800.  
[https://doi.org/10.1212/WNL.58.12.1791/SUPPL\\_FILE/THALTABLE.XLS](https://doi.org/10.1212/WNL.58.12.1791/SUPPL_FILE/THALTABLE.XLS)
- Thompson, P. M., Hayashi, K. M., De Zubicaray, G., Janke, A. L., Rose, S. E., Semple, J., Herman, D., Hong, M. S., Dittmer, S. S., Doddrell, D. M., & Toga, A. W. (2003). Dynamics of Gray Matter Loss in Alzheimer's Disease. *Journal of Neuroscience*, 23(3), 994–1005. <https://doi.org/10.1523/JNEUROSCI.23-03-00994.2003>
- Tran, J., Anastacio, H., & Bardy, C. (2020). Genetic predispositions of Parkinson's disease revealed in patient-derived brain cells. *Npj Parkinson's Disease*, 6(1).  
<https://doi.org/10.1038/S41531-020-0110-8>
- Valsky, D., Grosberg, S. H., Israel, Z., Boraud, T., Bergman, H., & Deffains, M. (2020). What is the true discharge rate and pattern of the striatal projection neurons in parkinson's

- disease and dystonia? *ELife*, 9, 1–27. <https://doi.org/10.7554/ELIFE.57445>
- Venkatraghavan, V., Bron, E. E., Niessen, W. J., & Klein, S. (2019). Disease progression timeline estimation for Alzheimer's disease using discriminative event based modeling. *NeuroImage*, 186, 518–532. <https://doi.org/10.1016/J.NEUROIMAGE.2018.11.024>
- Vercruyssen, S., Leunissen, I., Vervoort, G., Vandenberghe, W., Swinnen, S., & Nieuwboer, A. (2015). Microstructural changes in white matter associated with freezing of gait in Parkinson's disease. *Movement Disorders*, 30(4), 567–576. <https://doi.org/10.1002/MDS.26130>
- Vogel, J. W., Iturria-Medina, Y., Strandberg, O. T., Smith, R., Levitis, E., Evans, A. C., Hansson, O., Weiner, M., Aisen, P., Petersen, R., Jack, C. R., Jagust, W., Trojanowki, J. Q., Toga, A. W., Beckett, L., Green, R. C., Saykin, A. J., Morris, J., Shaw, L. M., ... Wollmer, P. (2020). Spread of pathological tau proteins through communicating neurons in human Alzheimer's disease. *Nature Communications* 2020 11:1, 11(1), 1–15. <https://doi.org/10.1038/s41467-020-15701-2>
- Vu, T. C., Nutt, J. G., & Holford, N. H. G. (2012). Progression of motor and nonmotor features of Parkinson's disease and their response to treatment. *British Journal of Clinical Pharmacology*, 74(2), 267–283. <https://doi.org/10.1111/J.1365-2125.2012.04192.X>
- Wang, Z.-T., Zhang, C., Wang, Y.-J., Dong, Q., Tan, L., & Yu, J.-T. (2020). Selective neuronal vulnerability in Alzheimer's disease. *Ageing Research Reviews*, 62, 101114. <https://doi.org/https://doi.org/10.1016/j.arr.2020.101114>
- Wennberg, A. M., Whitwell, J. L., Tosakulwong, N., Weigand, S. D., Murray, M. E., Machulda, M. M., Petrucelli, L., Mielke, M. M., Jack, C. R., Knopman, D. S., Parisi, J. E., Petersen, R. C., Dickson, D. W., & Josephs, K. A. (2019). The influence of tau, amyloid, alpha-synuclein, TDP-43, and vascular pathology in clinically normal elderly individuals. *Neurobiology of Aging*, 77, 26–36. <https://doi.org/10.1016/J.NEUROBIOLAGING.2019.01.008>
- Whitwell, J. L. (2010). Progression of atrophy in Alzheimer's disease and related disorders. *Neurotoxicity Research*, 18(3–4), 339–346. <https://doi.org/10.1007/S12640-010-9175-1/FIGURES/2>
- Wilk, A. J., Shalek, A. K., Holmes, S., & Blish, C. A. (2024). Comparative analysis of cell–cell communication at single-cell resolution. *Nature Biotechnology*, 42(3), 470–483. <https://doi.org/10.1038/s41587-023-01782-z>
- Wilson, D. M., Cookson, M. R., Van Den Bosch, L., Zetterberg, H., Holtzman, D. M., & Dewachter, I. (2023). Hallmarks of neurodegenerative diseases. *Cell*, 186(4), 693–714. <https://doi.org/https://doi.org/10.1016/j.cell.2022.12.032>
- Wilson, R. S., Yang, J., Yu, L., Leurgans, S. E., Capuano, A. W., Schneider, J. A., Bennett, D. A., & Boyle, P. A. (2019). Postmortem neurodegenerative markers and trajectories of decline in cognitive systems. *Neurology*, 92(8), E831–E840. <https://doi.org/10.1212/WNL.0000000000006949>
- Witt, S. H., Sommer, W. H., Hansson, A. C., Sticht, C., Rietschel, M., & Witt, C. C. (2013). Comparison of gene expression profiles in the blood, hippocampus and prefrontal cortex

- of rats. *In Silico Pharmacology*, 1(1), 15. <https://doi.org/10.1186/2193-9616-1-15>
- Wolfien, M., Galow, A.-M., Müller, P., Bartsch, M., Brunner, R. M., Goldammer, T., Wolkenhauer, O., Hoeflich, A., & David, R. (2020). Single-Nucleus Sequencing of an Entire Mammalian Heart: Cell Type Composition and Velocity. *Cells*, 9(2). <https://doi.org/10.3390/cells9020318>
- Wolters, F. J., Zonneveld, H. I., Hofman, A., Van Der Lugt, A., Koudstaal, P. J., Vernooij, M. W., & Ikram, M. A. (2017). Cerebral perfusion and the risk of dementia: A population-based study. *Circulation*, 136(8), 719–728. <https://doi.org/10.1161/CIRCULATIONAHA.117.027448/-/DC1>
- Young, A. L., Oxtoby, N. P., Daga, P., Cash, D. M., Fox, N. C., Ourselin, S., Schott, J. M., & Alexander, D. C. (2014). A data-driven model of biomarker changes in sporadic Alzheimer's disease. *Brain*, 137(9), 2564–2577. <https://doi.org/10.1093/BRAIN/AWU176>
- Zalewska, T., Pawelec, P., Ziabska, K., & Ziemka-Nalecz, M. (2023). Sexual Dimorphism in Neurodegenerative Diseases and in Brain Ischemia. In *Biomolecules* (Vol. 13, Issue 1). <https://doi.org/10.3390/biom13010026>
- Zeighami, Y., Ulla, M., Iturria-Medina, Y., Dadar, M., Zhang, Y., Larcher, K. M. H., Fonov, V., Evans, A. C., Collins, D. L., & Dagher, A. (2015). Network structure of brain atrophy in de novo parkinson's disease. *ELife*, 4(September 2015). <https://doi.org/10.7554/ELIFE.08440>
- Zhang, N., Gordon, M. L., & Goldberg, T. E. (2017). Cerebral blood flow measured by arterial spin labeling MRI at resting state in normal aging and Alzheimer's disease. *Neuroscience & Biobehavioral Reviews*, 72, 168–175. <https://doi.org/10.1016/J.NEUBIOREV.2016.11.023>
- Zheng, S. C., Stein-O'Brien, G., Boukas, L., Goff, L. A., & Hansen, K. D. (2023). Pumping the brakes on RNA velocity by understanding and interpreting RNA velocity estimates. *Genome Biology*, 24(1), 1–31. <https://doi.org/10.1186/S13059-023-03065-X/TABLES/1>
- Zhou, J., Gennatas, E. D., Kramer, J. H., Miller, B. L., & Seeley, W. W. (2012). Predicting Regional Neurodegeneration from the Healthy Brain Functional Connectome. *Neuron*, 73(6), 1216–1227. <https://doi.org/10.1016/j.neuron.2012.03.004>
- Zimmer, T. S., Orr, A. L., & Orr, A. G. (2024). Astrocytes in selective vulnerability to neurodegenerative disease. *Trends in Neurosciences*, 47(4), 289–302. <https://doi.org/10.1016/J.TINS.2024.02.008>